DOCUMENT RESUME

ED 188 442 FL 010 892

AUTHOR Pierce, Joe E.
TITLE The Nature of Natural Languages.
PUB DATE 79
NOTE 169p.
AVAILABLE FROM The Hapi Press, 512 SW Maplecrest Drive, Portland, OR 97219 ($7.95)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Acoustic Phonetics; Child Language; Consonants; Contrastive Linguistics; *Deep Structure; *Distinctive Features (Language); English; Form Classes (Languages); Japanese; Language Acquisition; Language Patterns; *Linguistic Theory; Morphemes; *Morphology (Languages); Phonemes; *Phonology; Speech Communication; Structural Analysis (Linguistics); Syntax; *Transformational Generative Grammar; Turkish; Vowels
IDENTIFIERS *Markov Processes

ABSTRACT

A variety of types of evidence are examined to help determine the true nature of "deep structure" and what, if any, implications this has for linguistic theory as well as culture theory generally. The evidence accumulated over the past century on the nature of phonetic and phonemic systems is briefly discussed, and the following areas of analysis are considered: (1) how phonemes are differentiated from each other by distinctive features and as fuzzy sets; (2) how morphemes come in two fundamentally different types, some of which form crisp sets and some of which are quite fuzzy; (3) how morphemes are combined into words which are not necessarily organized into classes in the same way that morphemes are; and (4) the fact that on the so-called syntactic level of analysis there are a variety of different types of units. Since it is proposed that the only thing that is transmitted from the speaker to listener is the vocal symbol, (i.e., the form), it is concluded that all linguistic analysis must proceed from the forms. It is suggested that at the point that phonemes are identified, morphemes are identified in terms of the sequence of phoneme combined with the statistical probabilities of occurrence. Then morphemes are identified on the basis of the phonemes and a Markov process (statistical probabilities) until a word is identified. Then the phrases are identified on the basis of the words and a statistical process and so on up the line. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED188442

THE NATURE OF NATURAL LANGUAGES

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Joe E. Pierce

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

FL010 892

JOE E. PIERCE

ISBN 0-913244-20-1
Library of Congress Number 79-88731

First Edition

Copyright, 1979
THE HAPI PRESS
512 SW Maplecrest Drive
Portland, Oregon, 97219
USA

TABLE OF CONTENTS

Introduction..... 1
Phonology..... 24
Morphology..... 49
First Language Learning..... 91
Syntax..... 109
Communication..... 131
Selected Bibliography..... 157
Index..... 160

I should like to dedicate this small book to all of those who labor to discover exactly how real languages work in real situations when real people are attempting to communicate real ideas with each other as we live our simple lives.

THE NATURE OF NATURAL LANGUAGES

The basic questions which any theory of language must attempt to answer are: What is Language? How do languages operate in conversation? What happens inside the brain of the speaker as he formulates a message and what happens inside the head of the listener as he hears an incoming stream of noise? These questions are crucial to understanding communication as well as language teaching and learning.

There are a multitude of answers to these questions presented by different linguists (c.f. the *Foundations of Linguistics* by Franklin C. Southworth and C. J. Daswani, Free Press, 1974 for a few of those in vogue at the present time). Much of the conflict within the field stems from sources which seem not to be understood generally.

Basically, much controversy appears to result from what different scholars believe to be occurring inside the brain of either the speaker or the listener. Transformational Grammarians are not at all hesitant to express this fact, but structuralists usually either deny it or sweep it under the rug. Essentially, anthropological linguists and structuralists both are attempting to understand what is going on inside the brain of a listener when he hears and attempts to decode a message, despite their protestations that they are only interested in behavior. Behavior is, in reality,

only their way of getting at the perceptions of the listener or speaker of a language. Such concepts as the phoneme can only have reality in the brain of the native speaker.

Few scholars today give serious thought to Edward Sapir's *The Psychological Reality of the Phoneme*, and yet this structural unit can only be relevant to linguistic analysis as a contrastive unit in the brain of the speaker or listener of a given language, i.e., as a psychological unit, whether or not one can define such a unit.

Phonemic analysis is an attempt to discern the systematic order in a sound system as perceived by a speaker. For example, a Turkish speaker of Turkish simply does not hear, and can be made to hear only with great difficulty, the difference between the vowel sound in the English word beet and the English word bit, because the range of his /i/ phoneme covers the ranges of both of the two English phonemes. The English speaker perceives this difference easily and automatically in normal speech without actually being aware that he makes the distinction until someone tells him so. Hence one says that the difference between /i/ and /ɨ/ is phonemic in English but the same sound difference is not phonemic in Turkish. Within the English speaking community, we have dialects in which the speakers cannot hear without specific instruction (Portland, Oregon), the difference between the words cot and caught, or, in fact, the difference between the phoneme /a/ and the so-called open-o, in any phonotactic position.

If one drops to the distinctive feature level, he is still faced with the same problems. We are talking about perception, that is, what goes on inside the brain of some speaker or listener of a language, whether we admit it or not. And is this not, indeed, "deep structure," whether we admit it or not. Hence, it becomes clear from their work that the structuralists are fundamentally concerned with the manner through which a stream of noise which impinges on the

ears of a human being is converted into a message. Transformational Grammarians, on the other hand, concern themselves basically with what goes on inside the brain of a person formulating a message to be transmitted. As a result of the above described situation, the problem becomes not "deep structure" or not "deep structure," but rather, what is the nature of the deep structure which we are all looking at and trying to understand by our study of languages.

The purpose of this book will be to attempt to look at a variety of types of evidence to see if we can better determine the true nature of "deep structure" and what, if any, implications this has for linguistic theory as well as culture theory generally.

Chomsky appears to use the label, deep structure, in two often conflicting ways, i.e., as (1) that basic wiring within the human brain given to all normal homo sapiens at birth which makes language possible, and I see no quarrel with this from any quarter, and (2) the specific way he says that the brain works, with which I think most linguists take issue in varying degrees. Hence, I shall talk about the "basic wiring" rather than "deep structure" and note that this corresponds to Chomsky's early discussions of deep structure but will turn out to be vastly different from his recent descriptions of deep structure.

All linguists would agree, I think, that the ability to handle vocal symbols is innately human. However, this inherent ability does not imply that a human society will automatically develop a system of vocal symbols any more than the innate ability to handle arithematic implies that a child will develop a complicated mathmatical system.

The fact that all human societies have language could have, based on what we know about the development and diffusion of cultural elements in general, three different implications. First,

all groups without languages could have been unable to defend themselves and were destroyed hundreds of thousands of years ago. Second, the idea or rudiments of language was "discovered" when man first became man, and all lanugages have a common root. Finally, since language is very old (probably over 3,000,000 years) and man presumably has a common origin from which he spread over the earth, language could have been invented by one group and spread by either diffusion or "stimulus" diffusion to all other groups. It is not even probable that all groups of human beings developed languages if they did not have it. Man is a remarkably uninventive creature, as shown clearly in the studies of diffusion of culture traits and trait complexes made over the past hundred years or so. Few inventions have been made twice in all human history. Virtually all of human culture has been invented only once and then spread around the world to be combined and recombined in a variety of forms. This, despite the fact that all of us are born with the innate ability to use these cultural inventions. There is little real doubt that language is not different from other cultural artifacts in the way it was created and spread around the world. Linguists face the same problems in dealing with language that all anthropologists face in dealing with culture. Well over 100 years of research into the nature of cultures around the world shows essentially that every logical possibility for any facet of culture exists somewhere on earth. This implies, logically, that whatever inherent wiring the brain has that makes culture (and language) possible is exceedingly flexible and non-restrictive. Language is a part of the "superorganic" level of nature (cf. A. L. Kroeber, *Culture; the Superorganic* in *The Nature of Culture*, University of Chicago Press, pp. 22-51, 1952) and is governed by the same basic nature as the rest of culture. Any theory of language must be compatible with this basic nature.

What was Kroeber's contentions, explicit and implicit, about the superorganic as different from

other aspects of reality studied by scientists? Essentially, Kroeber saw the world as beginning with the chemicals created when the universe was created, that is, inorganic substances. This, of course, generally coincides with scientific thought since the Greek Philosophers. Studying this aspect of the world has for many centuries resided with the hard sciences, such as, chemistry and physics, and the data from these sciences is characterized by being almost invariable, so far as our ability to measure things is concerned. The data here fit very well with the traditional Indo-European type of logic, which has been called the logic of the excluded middle, i.e., everything belongs either to the class "A" or to the class "Not-A." For example, an element is either oxygen or it is something else, and the characteristics which identify oxygen as a class of atoms are clear and immutable. This is only slightly over-simplified, but the reality of this level lends itself very well indeed to study through the logic of the excluded middle.

The next higher level, that is, the level of living organisms, is called the organic level of reality. Here the traditional type of logic of Indo-European speaking peoples is not so well suited to the study of the data, that is, it is not so easy to describe plants and animals as being members of either class "A" or class "Not-A." Such categories as species, families and even varities are not as clearly defined as are the elements on the inorganic level of nature. Further, on the lower level, measurements are usually taken in millimicrons or some other extremely small unit of measure. On the organic, on the other hand, measuring is done in feet, inches or even miles in some cases, such as grazing ranges or ecological niches. What all this means in simple English is that there is little variability on the inorganic level, but much more on the organic level. Also, while categories seem to never overlap on the inorganic level, there is considerable overlap on the organic level, e.g., the duckbilled

platypus is not absolutely a mammal nor is it absolutely a fowl, but it is possibly both, and yeasts and molds also have characteristics of both plants and animals.

What do the things said in the paragraph above have to do with the superorganic level of nature? Kroeber's contention, as I read him, was that just as the organic level came into being with the advent of life, the superorganic level came into being when, in the course of evolution, the human brain reached a size and complexity sufficient for man to operate on a symbolic level and create cultural systems. As best we can determine, this was probably between 2,500,000 and 3,500,000 years ago. On the inorganic level of nature the combinations of units and, indeed, all observable interaction is rigidly ordained by the nature of the universe. On the organic level such actions, reactions and interactions are still "limited" but not controlled by the nature of things. A species, for example, will have individuals in it that vary considerably in size, shape, color, and so forth, whereas, oxygen atoms are all alike, with, of course, such limited exceptions as heavy hydrogen. These are, however, few in number on the inorganic level. As a result of the differences between the inorganic and organic levels of nature, one needs different concepts, different theoretical tools, and, indeed, to some extent a new kind of logic to handle the organic level. These facts relate to the superorganic in that this new level, i.e., the superorganic, is not limited in any way that we have been able to discover. Perhaps there are natural limits of some kind, but in 1850 one of Tylor's avowed goals for the field of anthropology was the determination of what cultural elements were universal. From such information we were to be able to determine the "natural" man, that is, the basic wiring of the human brain that made culture (and language as a part of culture) possible. Over a hundred and twenty-nine years of research have turned up virtually nothing that is universal, and our ideas about the nature of man are now almost

limitless. Hence, in contrast with the two lower levels of nature, both of which have severe, but different, natural constraints and limitations imposed upon them by nature, the superorganic is limited only by the imagination of man, and so far as we have been able to determine to date, this is virtually limitless. This implies that totally new theoretical tools, concepts and logic are needed to handle the superorganic level of nature, but social scientists have for decades been trying to ape the physical sciences, because they have been so successful over the past hundred years or so. The reason that they were so successful lies in a combination of the logic which is an outgrowth of Indo-European thinking, which rests on the basic structure of proto-Indo-European grammar and the nature of the data on the superorganic level of nature. This logic works considerably less well on the biological, i.e., organic, level and still less well on the superorganic level, which is the level on which langauge operates and hence the relevance of this discussion to the subject matter of this book.

What exactly, then, is wrong with what linguists have been doing? This is exactly what we need to take a serious look at and do something about. In order to see how this applies, or does not apply, to linguistic analysis, we need to look at the traditional way Indo-European speakers view the world and what is wrong with this. We can, perhaps, best do this by looking at the work of an engineer, Professor L. A. Zadeh of the University of California. He was attempting to program a variety of types of language data to be fed into computers, and he soon concluded, along with many other people, that the problem was that the computer worked on the basis of what he calls "crisp" sets, whereas reality is actually a collection of "fuzzy" sets. The reason that it was necessary for him to create this dichotomy is that traditionally logicians, scientists and, indeed, linguists themselves, have only recognized

one type of set. This is not to say that arithmatic sets are not different from topological sets in certain mathmatical ways, but rather to say that all traditional mathmatical sets are "crisp." What, exactly, does this mean and what relevance does it have for linguistic theory? The label "crisp" to contrast with "fuzzy" sets had to be created in order for scholars to discuss a difference that had simply been ignored by previous generations. Taking topological sets as an example, a topological set is a collection of items in a topological space within a certain specified distance from a specified point. Other sets are defined in different ways.

The only point critical to an understanding of languages is that this space is bounded by a curved line which has no thickness. This means that since points in space also have no thickness, every point in any space being discussed is either inside the set or it is not, i.e., our point is either an "A" (a member of the set) or it is a member of the class "Not-A" (is not a member of the set). This is, of course, the law of the excluded middle.

It takes little imagination to note that at least a great deal of language, especially on the meaning level, consists of sets which do not fit this model, i.e., the edges of the set are not sharp, but rather are "fuzzy." Some things seem to be a member of one set at one time and not a member of that set at other times. Some things seem indeterminate as to which set they belong in, and so on. For example, what is a table? I think most of us have had the experience of referring to something in the real world as a "table-like thing," because it is not really a table, but it does not seem really to be anything else either, and the table is the most like the item we wish to talk about. The category in our culture which we label "table" is fuzzy.

Returning now to the differences between the inorganic, organic and superorganic levels of reality, one of the fundamental differences between

them is that the overwhelming number of classes (sets) of items on this level are crisp sets. On the organic level many, perhaps most sets are crisp with only a few being fuzzy, and these might be said to be relatively crisp or not very fuzzy. Now, on the superorganic level, the categories of a culture, which are symbolized by the verbal symbols that comprise a language, are practically all fuzzy. One of the basic weaknesses of virtually all linguistic theory proposed to date is the inherent assumption that linguistic categories can be treated as if they were crisp, when, indeed, they are not only fuzzy, but most of them are extremely fuzzy.

To illustrate this, let us take a brief look at Roman Jakobson's distinctive feature theory (cf. *Preliminaries to Speech Analysis*, Roman Jakobson, Gunnar M. Fant and Morris Halle, MIT, Research Report Number 13, 1952). Think briefly about Jakobson's early statements about the nature of a phonological system. He stated in the early 1950's that the phoneme was the basic phonological unit and that the variability of phonemes, which had been demonstrated again and again by linguists over the past hundred years, could be understood in terms of distinctive phonological features. This is related to our present discussion in that what he said about phonemes indicated clearly that he saw the phoneme as a crisp set, i.e., every noise produced by a speaker was a point in a class (set) or it was not. The listener supposedly could identify the phoneme in terms of the acoustic distinctive features (acoustic correlates) present in a stream of speech. This would be a classic representation of the crisp mathmatical theory of the type used by European scholars and logicians for the past 2,000 years or more. The fact that Jakobson and others have revised these arguments since then does not alter the basic problem, that is, the problem of seeing the world through crisp sets, which is the problem we are addressing at the present time.

Before proceeding, let us take a brief look at some of the evidence accumulated over the past century on the nature of phonetic and phonemic systems, because it is within this time frame that the acoustic spectrograph (cf. Martin Joos, *Acoustic Phonetics*, Language Supplement, 1948, for an explanation of the workings of this device) has confirmed many ideas that linguists had about the nature of human speech sounds and invalidated others.

First, let us take a look at the nature of consonants, their basic phonetic characteristics and the validity of distinctive features for explaining the way they function in the sound systems of natural languages. The easiest of these to deal with are, of course, the stops. Taking the English phoneme /p/ as a typical example of this type of consonant, we see that this sound is produced by stopping the flow of air out from the lungs by closing the lips. Hence it is called a bilabial stop. That is another way of saying that the articulatory distinctive features of the English /p/ are stoppage and bilabial articulation.

Since most scholars have recognized for a very long time that, while the articulation is important in learning to produce a sound and gives us a convenient way of talking about sounds, the articulation is not the thing that effects the human ear. The sounds produced by the various articulations are called acoustic correlates, and it is these acoustic correlates which impinge on the ear drum of a listener and which he must decode in order to understand the message that has been sent to him. Hence, these are what concern the linguist in his search to understand how languages work.

What then are the acoustic correlates of interest in discussing the English consonant /p/? Actually, a stop causes all frequences of sound above about 500 cycles per second to be filtered out of the spectrum of human speech, so the acoustic correlate of stoppage (as I shall refer to this distinctive feature) is the filtration of all frequencies above 500 c.p.s. from the speech spectrum. It is consider-

ably more difficult to state the exact nature of the acoustic correlate for bilabial articulation, but it seems clear that for English speakers at least these are different frequencies found in the spectrum of the releasë, that is, the puff of air (aspiration) following the release of the stop for /p/, /t/ and /k/, so for purposes of study, let us say that probably at some time in the near future we will be able to make statements, such as, bilabial articulation filters out none of the frequencies in the aspiration in the release and differs from alveolar and velar stops in that both such articulations do filter out part of the speech spectrum. Further, different concentrations of sound identify one stop as a /t/ and the other as a /k/. This may or may not prove to be true, but we can still use this to illustrate our point, because the truth of falsity of the statement does not alter the basic argument.

The relevance of what has been said above to linguistic theory, the superorganic level of reality and crisp vs fuzzy sets is that distinctive feature theory rests on the assumption that distinctive features are either present or absent, i.e., they serve the linguistic function of absolutely distinguishing one crisp set, phoneme, from another. If you remember, we said that crisp sets are sets of items in a space of some sort all of which are classed together because they either have or do not have a certain feature. They are labelled crisp, because everything in the universe either belongs to the set or does not belong to the set because if the distinguishing criteria is present, the item in question belongs to the set, if the item does not have that feature, then the item does not belong to the set. Hence, phonemes would be a clear example of traditional crisp sets if distinctive feature theory actually were true.

Do phonological systems actually work in this manner? We can, of course, view phonological systems in two slightly different ways, i.e., as consisting of bundles of features, phonemes, as

Jakobson did in the fifties, or as constellations of distinctive features which symbolize whole utterances, as most Transformational Grammarians appear to do in the late seventies. In either case, the distinctive features are pictured by the proponents of this theory as phonetic features which can only be "present or absent," i.e., they cannot be qualitative and partially present and partially absent. These, then, when present, signal to a listener that this or that phoneme or utterance was intended by the speaker. Both the phoneme and the utterance in this sense, then, are crisp sets which are uniquely identified by the presence of the specific distinctive features.

Since our problem, i.e., "how does a linguistic sound system work?" is the same in either case, and the phoneme is more easily discussed, let us look a little further into the possibility that these hypothetical sound units are crisp sets identified by distinctive phonetic features which are acoustic correlates of articulatory movements. We saw for the English consonant /p/ that this idea worked fairly well, and /p/ is quite representative of all stops and affricates. Actually, the arguments given to support the theory of distinctive features for stops are true for fricatives, nasals and semivowels, as well, but liquids and laterals pose some special problems. However, the real problems arise when we attempt to handle vocalic systems as crisp sets. The system then would become a system of crisp sets each identified by a unique constellation or bundle of distinctive features.

In order to understand the nature of phonological systems, as systems, we need first to look at the phonetic characteristics of examples from real speech (cf. Joe E. Pierce, *A Spectrographic Analysis of English Vowel Phonemes under Primary Stress*, Linguistics, Volume 84, pp. 41-84, 1972 for more details). There have been two somewhat contradictory theories in vogue as to the nature of the phoneme for quite some time (contrast Kenneth L. Pike, *Phonemics*, University of Michigan Press, 1947, with

Jakobson, Fant and Halle, op. cit. for examples of these two divergent approaches). Though Pike never used (and probably never heard) the label fuzzy set, the picture that he paints of a phoneme is clearly that which Professor Zadeh attempts to explain with his fuzzy set theory of mathmatics. The phoneme is pictured as a "range" of sound by Pike with each of these ranges being conditioned or altered by both "free" and "positional" variation. There is, at least in theory, a core area of sound which represents the "central tendency" of each phoneme and this has a fuzzy edge in which the actual sounds represent different phonemes dependent upon a variety of causes. In this view of phonology we see the basic characteristic of the phoneme varying from utterance to utterance. Jakobson's view, on the other hand, is that of a phoneme seen as a traditional crisp set. A phoneme is a bundle of sound features, some of these are distinctive and some are non-distinctive. If one changes the distinctive features, he gets a different phoneme, and as a result a different utterance. Hence, the logical implication of this theory is that the features which vary are the non-distinctive features with the one exception that "conditioned" or "positional" variation can involve distinctive features. Our question at this point is, "do vowel sounds operate according to the general view common prior to Jakobson's publication of distinctive feature theory, as exemplified in the writings of Kenneth L. Pike and others, or do they operate as described by Jakobson and his followers?

There is evidence from a variety of sources and of a variety of types to indicate the "real" nature of the sound signals we are dealing with. Here, I commit myself to being a "God's Truth" linguist, because I do, absolutely, believe that the language we want to describe exists "out there" in reality, just as does a tree (which a botanist attempts to describe), and our job as linguists is to describe the nature and characteristics of the

language we wish to discuss as accurately as possible. It does not, indeed, matter whether one achieves the description while sifting through data or "while shaving" to quote Emmon Bach, but I do not see how one can hope to be accurate, except through careful study of real speech sounds. With this prejudice (or theoretical bias, whichever?) in mind, perhaps my line of reasoning can more readily be followed.

Returning, momentarily, to 1952 and the publication of Jakobson, Fant and Halle's *Preliminaries to Speech Analysis*, I saw this historic publication as the savior of linguistics. It was several years after its publication that I began to have some reservations about its reliability as a tool for describing natural languages, and it was not until the early 1970's that I began to have serious doubts. Hence, I did not approach the theory with antagonism and attempt to disprove it, but with the assumption that it was correct and that its basic tenets could be demonstrated. As a result, the fact that after many years I propose that real languages operate in the manner to be proposed in this book reflects a slow and reluctant withdrawal from my earlier theoretical views (cf. Joe E. Pierce, *A Theory of Language, Culture and Human Behavior*, The HaPi Press, 1972 for an explanation of my beliefs at that time). It also seems to me that this sequence means that I "discovered" something about the way languages really work in the process.

My first doubts appeared with the publications in the late 1950's of the Haskin's Laboratories. In very simple English what Haskin's Laboratories did was to segment isolated words and remove the vowel nucleus from a given word, for example, the vowel sound in the word beet, i.e., /i/. Then they would take this vowel sound and transfer it to another word and insert that vowel for the vowel that was supposed to be there, for example they would take the /i/ vowel removed from beet and insert it into the word pack, i.e., the vowel sound in pack is /æ/. Now, one would expect the resultant word,

i.e., pack with the new vowel /i/ in it instead of the /æ/, to become peek, but it did not. Instead, such substitutions often resulted in a native speaker of English identifying the new word as poke or something else equally unlikely from the point of view of traditional linguistics. If distinctive feature theory were valid, such experimental results would not be possible. My first reaction to such studies was that there was something wrong with the research. The first possible flaw, and what seemed to be the most probable, to come to mind was that the phonemic cuts were not made accurately. None of my worries proved to be true, however, and in the end, I had to admit that the sound in the word peek, which most, if not all, speakers of English would identify as an /i/ was phonetically vastly different from the vowel sound in the word peep, which most English speakers would also identify as an /i/. Further, if these syllabic nuclei were to be identified by native speakers on the basis of distinctive features, they should have been identified as the same, even if they were vastly different so far as the non-distinctive features were concerned. In order to attempt to get to the bottom of this problem, I designed an experimental study of English vowel phonemes under primary stress (op. cit., 1972). I had made three extensive attempts to identify the distinctive features of some of the English syllabic nuclei (vowel phonemes if you wish); one in 1952 with Robert Stockwell (unpublished) under the direction of Martin Joos at the Summer Institute of Linguistics as Michigan University, a second one which was somewhat inconclusive but which resulted in the publication *Phonemic Theory and the Analysis of English Syllabic Nuclei*, Linguistics, Volume 7, pp. 63-82, 1964, and a somewhat more refined one which remains unpublished, still with no notable success in discovering the distinctive features of vowels or syllabic nuclei. In each of the three above cases I concluded at the end of the experiment that the problem with the

research design had been too many uncontrollable variables. In order to make this one last attempt as meaningful as possible, I used a single informant, and took all sample measurements from the syllabic nuclei of monosyllables under primary stress. This eliminated the possibility that the word uttered could be identified on the basis of anything other than the sound itself. I reasoned that distinctive features would have to be identifiable under such rigidly controlled conditions, and I would be able to find them easily.

What were the results of this latter study (op. cit., 1972)? The actual picture which emerges is considerably more complex than either of the two previously described theories would indicate, i.e., the theories of Pike or Jakobson. We see not only a range of vocalic qualities with clear phonemic overlap, i.e., the same sound representing a phoneme /i/ in one instance and a phoneme /ɪ/ in another. We also see sometimes radically different ranges of sound representing what is identified by the native speaker as a given syllabic nuclei when in different phonemic environments. We not only see that each syllabic nucleus is fuzzy (and clearly not crisp), but in several different environments the fuzzy sets which represent a given nucleus, e.g., the vowel in beet, are quite different. This agrees perfectly with the findings of the Haskins Laboratory research. The reason Haskins Laboratories got the results they did is the fact that English vocalic nuclei are not crips sets identified by the acoustic correlates of articulations which are perceived as distinctive phonetic features but rather are ranges of sound, clearly fuzzy, and since the range for each nucleus is different in different environments, quite naturally the sounds heard were identified as a totally different sound when removed from the environment in which it was originally produced and placed in a different one.

The results described above should have been clearly predictable from our observations of

spectrograms and other acoustic data even as early as 1951, and it is my judgement that the only reason we did not is our Indo-European tradition of attempting to deal with everything in terms of crisp sets. What do spectrograms tell us about human speech. First, as the tongue position goes up, formant number one goes down, and as the tongue goes down, the same formant goes up. Anyone who questions this should consult my article, *Spectrographic Study of Vocalic Nuclei*, Language Learning, Volume XII, Number 3, 1963, pp 241-47 and replicate what was done there to verify that what we see here is more than just an accidental parallel with the traditional vowel charts drawn up on the basis of articulation. Further, as the high point of the tongue is moved from front to back the position of formant number two moves down, and, of course, as it moves forward the formant moves back up. This conforms to Martin Joos two-formant theory of vowels (cf. Martin Joos, op. cit., 1948), except that one can also position the second formant by rounding the lips. The problem is that since one can move his tongue or lips in a continuous motion, there is not even the logical possibility of generating distinctive features comparable with the stoppage or bilabial articulation as described earlier with consonants and their distinctive features.

The problem with theoreticians in any field, and especially in linguistics, is that they get so carried away with theorizing that they seem at times to forget what they are theorizing about in the real world (cf. Paul L. Garvin's review of the *Premilinaries to Speech Analysis*, Language, Volume 29, Number 4, 1953, pp 472-81 for several examples of the nature of such theorizing). I find it necessary to agree with Saussure's dichotomy of *Lanque* and *Parole* (cf. Rulon S. Wells, *Sassure's System of Linguistics*. Word, Volume 3, Number 1, 1947, pp 15-18 for explanation). However, I am not sure that I see the dichotomy in exactly the

same light as many other linguists do. As I see it, langue or language, is a hypothetical system which reflects a classification of the functional units in the speech of a given group of speakers, i.e., to use Sturtevant's labels, a speech community. Since we are currently talking about the sound systems of languages, we are here concerned with the phonemes or the functional sound contrasts for a language. The phonemes exist nowhere but in the heads of native speakers of a language, and since no two speakers appear to have exactly the same system in their heads, a phonemic system for a given language must represent some sort of statistical average of the systems in the heads of a number of speakers in that language. Hence any discussion of language structure must involve statistical linguistics, not merely logical hypotheses about bipolar contrasts. Further, the statistics must somehow incorporate phonetic data, i.e., parole or speech. We shall discuss later in some detail the problems with such concepts as linguistic competence.

What I hope you have derived from what has been said above is that in the real world of human languages there appears to be two very different types of phonemes, one of which is typical of the crisp sets utilized by Indo-European logicians and scholars in other scientific fields for many generations and the other one represented best by what Zadeh has described as fuzzy sets (cf. L. A. Zadeh, *Fuzzy Languages and their Relation to Human and Machine Intelligence*, Electronic Research Laboratories, University of California, Berkeley, Publication ERL-M302, 1971 for a clear exposition of the nature of fuzzy sets and a bibliography). In a way this is no revelation to anyone, as we have classified speech sounds into consonants and vowels for centuries, but one of the basic problems has been that it has been impossible to define a vowel or a consonant in terms of any real phonetic criteria. They are probably best defined in terms of distribution or function as Charles

Hockett did in his, *A Manual of Phonology*, Indiana University Press, 1955. The central point here would be that there is no acoustic distinctive feature that could be labelled as "consonontal" nor indeed one that could be labelled as "vocalic," and any theory of phonology resting on such an assumption must be discarded (for a simplified exposition of this theory cf. *Foundations of Linguistics*, Franklin C. Southworth and C. J. Doswani, The Free Press, 1974, pp 72-76).

In order to formulate a theory for language structures, we need to know very precisely how languages work. How are we going to find out how languages work? The only way I can imagine to do this is to describe the actual workings of hundreds of languages and then from these descriptions, extract the common core. This is what descriptive linguistics has been all about since its founding in the late 19th century. How does one, then, go about describing languages? In earlier ages we saw scholars, armed with the belief that Latin was the prototype for language, going about the world describing languages as distorted Latin. Most of what some linguists call "linguistic universals" today are a product of this method of describing languages and not really a reflection of similarities in the way languages work. If scholars assume before they look, on whatever grounds, that all languages have nouns, then later scholars will find nouns in all languages whether such a class really exists as a functioning part of the structure of that language or not. This is true because of the earlier assumption and tells us nothing about how that language really works. If we are to succeed in our quest for linguistic universals, then our descriptions must assume as little as is possible when we make our original description. The distortion of English grammar which resulted from earlier grammarians asking the question, "what are the nouns in this language?" instead of asking the relevant question, "waht are the parts of speech

(word classes) in this language?" still plagues us today as we attempt to teach English (for further discussion of this problem, cf. Joe E. Pierce, *How English Really Works*, The HaPi Press, 1979, pp i-viii). We must have accurate, objective descriptions of how languages operate as systems, and these descriptions must be derived from a careful study (unbiased by unrealistic assumptions) which attempts to "discover" the structure that comes naturally out of observations of the behavior of speakers of the language.

How do we derive such systematic descriptions? Quite simply by studying very carefully a sample of the language. This brings us to one of the bug-a-boos of linguistics today, i.e., the closed corpus study. The structuralists have been severly criticized by some scholars for utilizing a closed corpus procedure. Some have argued that when you look at a closed corpus, you do not really know what you might find outside the corpus, and what good is the description of such a corpus anyway. Such arguments bring us directly to generative grammar. Despite claims to the contrary by some critics, I have never seen a grammar that did not claim to be generative, at least in that the patterns found within the closed corpus could generate, if not all possible sentences, at least the overwhelming majority of such sentences. All grammars are generative and always have been, so I hope we can forget about that bit of folklore and get on with the serious study of languages.

The real question which we must face is, how do we reach an understanding of the structure of a given language so that later we can understand the basic nature of languages in general by comparing descriptions of many languages and from this "discover" the nature of the basic wiring in the human brain that makes human speech possible. This absolutely MUST be discovered because there is an infinite number of possible guesses, such as that made by Transformational Grammarians, any one of which might be just as good

or as bad as any other. The problem with all of them is that the assumptions made by the scholars who create the systems of description usually predetermine the universals which will be derived from studies made under such theories, and as a result the view of the human brain which we project from our research reflects the ideas we started with whether or not these reflect anything in the reality that is the language.

One often hears the criticism that the mechanical study of a closed corpus has had little success. As I see these statements, they stem from two basic sources. (1) The writers fail completely to understand what a structural linguist does when he studies a closed corpus, draws his formula (for generating utterances) from that corpus and produces a grammatical description. (2) if the results you get do not agree with the writer's "intuition" about the language, then your study is said to have yielded little of value. In response to several articles I had read in 1966-67, I conducted a small study and wrote the article *A Look at the So-called -Ing-forms of English Verbs*, Linguistics, Volume 50, 1970, pp 59-69, simply to demonstrate exactly what one should do if he wished to "discover" the true nature of the workings of the grammatical structure of a language. My reasons for doing this was that all of the things that were being said in the literature, mostly by Transformational Grammarians, showed clearly that they not only did not understand what structuralists were doing, they were determined not to give the ideas any real consideration. The amazing thing about the little study was that I "discovered" some things about English grammar that amazed me. I would have adamantly denied that these things were true, had someone asked me before I did the study. This proves, at least to me, that one's intuitions about his native language reflects what he has been taught about the language far more than what is true about the way the language works, even if he is a trained linguist. The most striking example

is that under no circumstances can an -ing-form, such as going, function as a sentence predicate in English. It is precisely this fact that forces us to utilize so-called "helping verbs" in many sentences. These helping verbs drive foreign learners up walls, because they cannot imagine what the forms are doing in the sentence. They seemingly are meaningless to such students. As a result they consistently leave them out. They must be there because English sentences have to have predicates and the -ing-forms cannot function as a predicate. This really needs to be taught, especially to foreign students, and it would not hurt native speakers to know it either.

The point is very clear, then, that what we "intuitively" believe to be true about our native language coincides with what we were taught in school, and what we were taught in school is a direct by-product of the Latin Grammatical tradition. This tradition so warps and distorts the structure of English, and I suspect many other languages, that the grammar is not recognizable to anyone who speaks the language and who has not been put through the propaganda mill in the school systems of this country. As a result of these facts, "discovery" procedures are absolutely essential if we ever hope to know anything about the way this language really works.

None of the criticisms that I have seen levelled at the closed corpus technique for discovering the structure of languages are valid, and this technique has some very definite advantages. However, to understand the advantages one must know how to apply the technique properly. First, any study of a language involves a corpus of data, even if it is all that the linguist knows about his native language, and if the linguist is a mature adult, say thirty years of age or older, then it is, for all practical purposes closed. The weaknesses of dealing intuitively with this type of corpus is that it is too large to handle carefully and the linguist is too limited in the things that happen

to come to his mind. On the other hand, if he takes, for example, a five percent sample of a book, that is, every twentieth page, and the book is eighty thousand words long, one is very likely to find most anything that is significant about that form somewhere in the sample. The best way to really discover the significant grammatical features of any language is to take some one form, such as, the -ing-forms of English verbs, and extract every occurrence of that form in your sample. Then you force yourself to study carefully each occurrence of that form. In the end, your explanation would then cover every actual usage of the form found. The problem with most linguistic research is that the scholars treat what they think will make the strongest case for whatever they believed to be true when they started, and that is no way to reach an understanding of any natural phenomena. At least with a closed corpus procedure one is likely to cover most of the range of usages for specific forms and in about the frequency they would be used generally in the particular style of the language that is being studied. This would most certainly yield a picture of the grammar which more nearly reflected, accurately, the structure of the language than would any intuitive study, because you do not allow yourself to sweep facts (such as, the past tense in English means past time less than half of the time in real sentences) under the table and ignore them. You force yourself to adequately treat every example of the form found in the data. The results of such studies usually surprise the linguist as much as they surprise anyone else, and it is precisely this fact that makes the procedure invaluable to the science of linguistics.

Having discussed briefly in this introduction some of the various ideas about how languages work and possible weaknesses in them, let us now begin in the next chapter to look at linguistic structures in a somewhat different light.

CHAPTER II

Linguists, whether expressly or implicitly, have all looked at language in terms of levels of analysis. Each of these levels has been treated as a separate system, and the significant difference among linguistic schools of thought has been exactly which levels and how many levels they recognized.

Let us now take a brief look at the historical evolutionary sequence which lead up to the present situation. Since 1920 the field has been rocked by at least two major revolutions. The first was the structuralists revolution, led in Europe by Ferdinand de Saussure and Prince N. S. Trubetskoy (cf. N. S. Trubetskoy, *Grundzuge der Phonologie*, Cercle Linguistic de Prague, 1939 and Ferdinand de Saussure, *A Course in General Linguistics*, translated by Wade Baskin, Phisolophical Library, New York, 1959), and in America by Leonard Bloomfield (cf. *Language*, Holt, Rinehart and Winston, 1933, and Edward Sapir, *Language*, Harcourt, Brace and World, Inc., 1921). The linguistic schools represented by these scholars were discovering many of the new concepts in structuralizing at about the same time, one group in Europe and one in America. There was perhaps some interaction between the two, but for the most part the discoveries appear to have been only slightly influenced by studies on opposite sides of the Atlantic.

Looking historically at the evolution of linguistic theory which led up to the formation of the two schools mentioned in the paragraphs above, we see that the ancient Greeks used the concept of grammar to encompass every aspect of language study from phonology up through grammar and syntax as well as stylistics, semantics, and so on. By 1850 most linguists had begun to look at languages through one of two different lenses, i.e., phonology or grammar, both considered as if they were more or

less independent systems. The grammatical level included all types of language study except the sounds until near 1900. Shortly after the turn of the century linguists began to realize that there was a difference between phonetics and phonemics (cf. Edward Sapir, op. cit., 1921). This recognition of a difference led to a break into two different ways of looking at the sounds of language. At about the same time one sees the break up of grammar into the two level, morphology and syntax (cf. any university catalogue in which linguistics is a part of the curriculum in the early twentieth century). This division represented a recognition that the morpheme, i.e., the basic unit of grammatical structure of a language, represented a fundamentally different type of unit from those on the "higher" levels of analysis. With the emergence of logical empiricism - logical pragmatism as the dominant philosophy at the time, linguistic scholars were striving to achieve a "scientific" way of handling linguistic data. The morphological level of analysis lent itself to the development of rigorous scientific techniques, whereas syntax did not appear to. This emphasis on the scientific method, along with the influence of Leonard Bloomfield, with his orientation toward a behavioristic theory of language, led the structurnalists of the 1930's and 40's to devote an inordinate amount of time and effort to studying the phonemic and morphological levels of analysis, leaving phonetics and syntax pretty much alone. This neglect on the part of the structuralists led to the second great revolution, i.e., the rise of transformational grammar as an alternative (cf. Naom Chomsky, *Syntactic Structures*, Mouton, 1957 for the beginnings of this movement). Fundamentally, the transformational revolution was a revolt against the structuralists who had omitted what to many represented the fundamental part of language study, i.e., the meaningful sequences produced in normal conversation or writing. Most of the work of transformational grammarians has been

centered around syntax with this school studying two levels of analysis, that of the distinctive feature and that of the sentence.

If one looks at the levels of analysis mentioned in the paragraph above, he will note that phonetics is at the bottom. This was defined out of linguistics in the thirties, because it represented unstructured data and linguistics is interested in the structured aspect of human speech. This, however, still left us with two phonological levels, because Jakobson, Fant and Halle (op. cit., 1952) introduced the distinctive feature level. The introduction of these units left linguistics with two ways to go, the utterance can be considered to be a unique constellation of distinctive features, a la some of the transformational grammarians, or as a sequence of phonemes with each phoneme being a unique bundle of such distinctive features. Different groups of linguists have, since the mid-fifties, chosen to go either of these two ways. However, in this book, we are going to take a slightly different track and once more reduce the number of levels in the phonological domain to one, and that one will encompass some of distinctive feature theory and some fuzzy set theory, because we believe that this is a much better, clearer reflection of what actually goes on in natural languages. That is, the discovery of the phoneme was an important theoretical break through for linguistics and so was the discovery of the distinctive feature, but the discovery of fuzzy sets gets us one step closer to the reality of natural language structures.

Before going into a study of the sound systems of languages it is necessary to understand a little more about how meaning is transmitted, because this has been one of the major stumbling blocks in the study of languages over the past several decades. Most linguists recognize that part of the meaning of an utterance is carried by the social situation. That is, if one is at a race track and mangles the word "bet" so that it is very unclear phonologically, it will be understood because there

is no reasonable alternative for the listener to pick up. However, in a real situation, it is quite possible for a worker in one of the stables to be talking about making his "bed," in which case the listener may completely misunderstand the speaker even though his articulation of all of the distinctive features of sound are very clear. Further, most linguists realize that much meaning is carried through gestures, and so we talk about kinesics. One has to be very careful when handling utterances to know exactly what is carried by the stream of noise and what is carried by some other elements, such as hand signals. Paralanguage is also usually noted by linguists. The voice quality can carry a great deal of meaning, and one must not attribute this to the symbols in the sentences uttered. There is also the very important fact of the life histories of the individuals involved in the conversation (cf. Joe E. Pierce, *Life Histories of Individuals and their Impact on International Communication*, in the book edited by Fred L. Casmir, *Intercultural and International Communication*, University Press of America, 1978, pp 522-540 for a further discussion of this point). The fact that human speech is redundant is recognized by most scholars (cf. Southworth and Daswani, op. cit., 1974, pp 90-91 for a simple illustration of this). However, the Markov process is exceedingly important, despite the protestations by some linguists that it is not. The explanations and refutations in most books concerning this process in human speech illustrate that the writers do not understand the process and not that the process is not operating. The very fact that one can leave blank spaces in a sentence and have a native speaker fill in the blanks correctly demonstrates the usefulness of this concept in linguistic theory. Further, a number of research projects by Bell Telephone Laboratories and Haskins laboratories are explanable, so far as I can determine, only in terms of Markov solutions. We will get back to this later, as we discuss various items within our framework for linguistic theory.

Let us now take a brief look at the evidence which supports a distinctive feature theory. Our purpose will not be to prove or disprove the basic theory but rather to see just how far it can be found to be valid. There can be no question as to the validity of portions of the theory and no question as to the importance of the contribution of Jakobson, Fant and Halle to the theory of language study. Jakobson generally follows the premise that all aspects of the phonological system can be understood in terms of bipolar contrasts (cf. Paul Garvin, op. cit., 1953 for an explanation of this). He establishes (Roman Jakobson and Morris Halle, *Fundamentals of Language*, 1956) twelve bipolar contrasts. Generally, he seems to hold that all languages differ only in the way that they utilize these twelve distinctive feature contrasts. These are:

1. Vocalic-nonvocalic
2. Consonantal-nonconsonantal
3. Compact-diffuse
4. Tense-lax
5. Voiced-voiceless
6. Nasal-nonnasal
7. Discontinuous-continuant
8. Strident-mellow
9. Checked-unchecked
10. Grave-acute
11. Flat-plain
12. Sharp-plain.

As we stated in chapter I there is absolutely no valid phonetic correlates for Vocalic-nonvocalic and Consonantal-nonconsonantal, because the same sound can be both a consonant and a vowel, even in the same language (cf. Charles Hockett, op. cit., 1955, and his coining the terms "omnipotent" phonemes and "demivowels" to cover exactly the problem that there can be found no articulatory or acoustic distinctive features for vocalic and consonontal). Compact-diffuse and Tense-lax are at best the opposite ends of a continuum and

hence do not lend themselves very well to a plus vs minus type of logical application. The next three, i.e., voiced-voiceless, nasal-nonnasal, and discontinuous-continuous, are virtually letter perfect for distinctive feature type analysis. All the rest are suspect on the grounds that they are opposite ends of a continuum and as a result they are subject to infinite gradations. Hence they do not make really good distinctive features.

If the distinctive features are, indeed, gradations along a continuum and not present vs absent phenomena, then you have accomplished nothing in going from phonemes to distinctive features. You simply have to face the same problems on a lower level. Let us take a look, from a phonological point of view, at a possible English utterance and see if we can imagine how a listener decodes the message. First, take the utterance:

/hay/.

I purposely leave off the spelling, because there are several different meanings for this utterance, given different situations, and one of the points that must be made is that all that is transmitted by a speaker and all that is heard by a listener is a symbol, nothing more. Not only is the only thing transmitted a symbol, but it is only a phonological symbol, so different writings mean nothing in the spoken language. If the speaker is looking straight up into the air, and I can see what he is looking at, I will interpret the word to mean high, as high up, but if he is looking at me, smiling, and I have not seem him before for some time, I will interpret the word as "hi." To my knowledge no linguist has ever denied these facts, so you may well ask, why bring that simple thing up. The reason is very simple, much of what I see in the literature about what this or that symbol means or this or that sequence of symbols mean have incorporated in them much that cannot come from anything but the social situation or something else outside the linguistic system. The real point

is that language is a system of symbols, nothing more, and all too many linguists try to make it something more.

If we dissect the utterance /hay/ into its component phonetic elements, we find an initial puff of air, which we can quite legitimately call aspiration and treat as a distinctive feature. It can be so treated, because if that aspiration is not there, there is a clear glottal stop. It is really the stoppage of the glottis which distinguishes /ay,/ that is, I from /hay,/ that is, either high of hi. Since we do not utilize the glottal stop as a phoneme in English, most linguists would be satisfied to call the aspiration a distinctive feature, and the only distinctive feature that told the listener that the word high was intended and not the word I, from a phonetic point of view as well as from a phonemic point of view. What this illustrates is that we can distinguish the symbols, at least to some extent on the basis of distinctive feature type phonetic data, but we cannot tell the difference between words that sound the same or morphemes that sound the same on this basis. The sounds /ay/ can be interpreted as a glide from the low central or low front position up toward the high front position of the tongue, and it is possible to consider this gliding as a distinctive feature, but it is still qualitive and not the ideal type of present vs absent phenomenon we would like.

The real question, and one that is often raised in linguistic discussions, is exactly how one distinguishes the meanings of different utterances. Actually, to understand how language really works, one has to realize that the only meaning that exists in a speech situation comes from inside the head of the listener. You do not have meaning transmitted from person to perons, only symbols. Once a listener has heard a stream of noise and identified a symbol, hopefully, correctly, but not necessarily correctly, he then reaches into his own brain and pulls out

whatever meaning he has stored opposite that word, given the sequence he has identified. Hence, the problems of linguistics become the problem of finding out how we identify symbols, and the meaning is actually irrelevant. The meaning that a given individual will apply to any sequence is dependent on his experience with that sequence. This has been demonstrated many times by linguistic research. Go out into the community and ask a hundred people what anything you want means, and you will find that you are very likely to get one hundred different answers. We will return to this point later, when we begin to deal with morphemes, because that is where the meaning comes into play, but it is important to realize that what the sender intends may be very very different from what the listener hears.

Returning now to the concept of distinctive features, let us take a look at the consonants. If we begin with what is the normal upper left hand corner of a consonant chart, we find stop consonants, and clearly stoppage is an ideal example of our theoretical distinctive features. It seems quite reasonable to assume that the points of articulation can be treated as if they were, so far as the identification of dental and bilabial articulation are concerned, but this is not necessarily true for velar, palatial and back velar. In fact, it would appear to be difficult if not impossible to so treat these consonants. Further, the next type of articulation down from the top is fricative, and while it is theoretically possible that this works as a distinctive feature in the bipolar contrastive sense, it is far from clear either that they do or what the distinctive acoustic correlates to such articulations might be. However, let us be generous, since we have no clear evidence to the contrary, and say, maybe, to the question, are these really distinctive features? Nasals clearly are, as is lateral articulation. I can point to clearly identifiable articulatory pat-

terns on spectrograms when lateral articulation or nasals are present in a stream of speech. The so-called semi-vowels or semi-consonants /w/ and /y/ can easily be treated as glides which clearly differentiate them from phonetically similar vowels. Hence we see that most of the consonant system can easily be treated as a system of fairly rigid phonemic units, distinguished from each other by identifiable distinctive features in the acoustic patterns of human speech. It is not clear that this is actually the way that human beings operate, but it is quite possible, and that being the case, let us postulate at this point that they do. We still have the liquid phoneme /r/ in English and some similar phonemes in other languages which quite probably do not operate on this basis. It is also quite possible that /l/ does not, and we shall see more about that in a moment. At least it is very clear that some phonemes do operate exactly as Jakobson said they should, but we are not sure exactly how many and which ones do, though we can point to the stops and probably the fricatives and nasals as quite sure.

In order to better understand the nature of vocalic nuclei, i.e., vowels, we need to take a careful look at the spectrographic patterns left by such sounds. In Chart I, pictured on page 33, you will see a schematic diagram of the results of a large sample, about 1700 vocalic nuclei, made by a single informant, of English vowels in monosyllables under primary stress. This was done in order to reduce the uncontrollable variables so that we could be absolutely sure that it was the vowel and the vowel alone that was distinguishing the morphemes uttered. By using a single speaker and monosyllables under primary stress, we eliminated many variables which may not be significant in normal speech but which might be clearly visible in the spectrum. We can easily see that the patterns of these vowels plotted after Joos (op. cit., 1948), i.e., the first formant plotted on a vertical axis and the

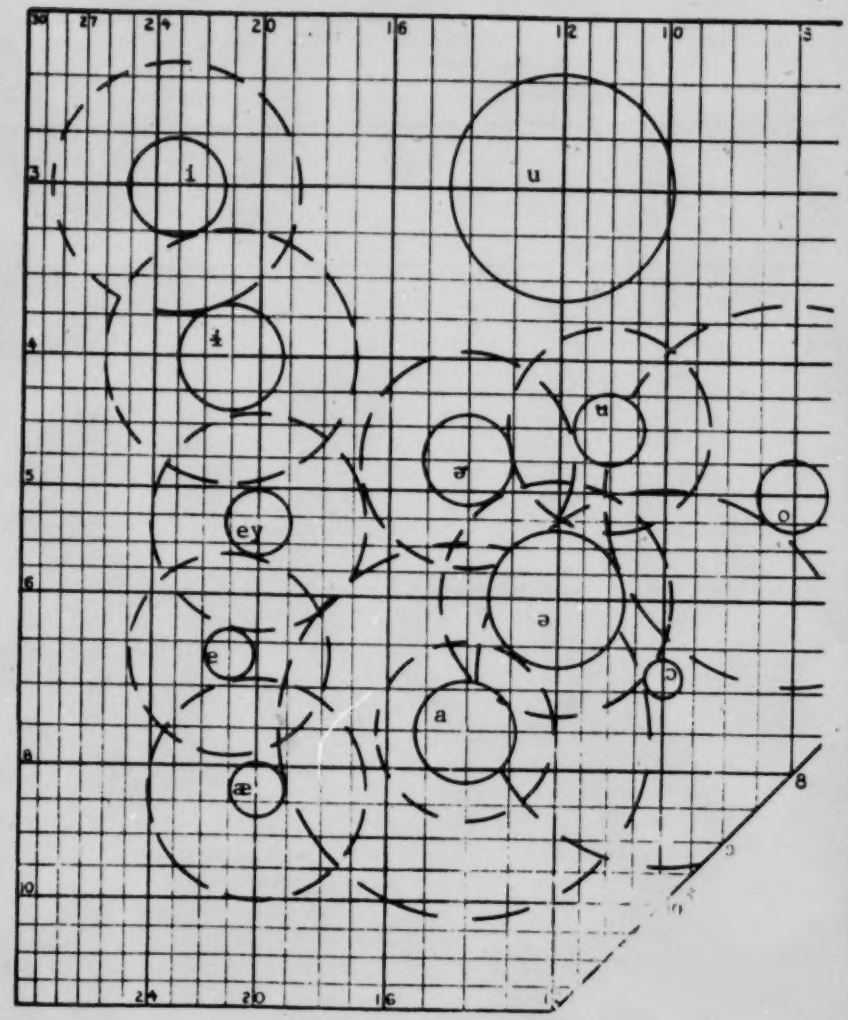


CHART I

Above we have a schematic diagram showing the nature of the distribution of approximately 1700 English vocalic nuclei, produced by a single speaker in single morpheme utterances.

second formant plotted on the horizontal axis with the point of origin being the upper right hand corner rather than the lower left hand corner (this is done because the first formant goes up as the jaw or tongue position is lowered and goes down as the jaw or tongue position is raised, so by inverting the picture we get a reflection of the articulatory movement) gives a pattern reminiscent of the picture of a phoneme described by structuralists during the forties and fifties, i.e., as a scatter pattern with a clear central tendency and fuzzy edges.

What has been said above could possibly be explained by distinctive feature theory, given enough assumptions, but let us for the moment attempt to step outside our Indo European philosophical tradition of logic and see if we cannot find another, perhaps more accurate, way of explaining what we see. The full body of evidence on English vocalic nuclei shows clearly that phonemic overlap (i.e., the same sound representing different phonemes at different times) really exists. In the overlapping of the ranges plotted for the vocalic nuclei /i/ and /ɨ/, that is, the nuclei found in the words <u>beet</u> and <u>bit</u> respectively, we find that in some cases the sounds were, indeed, not the same, because the /ɨ/ would be gliding down and back, but the /i/ would be gliding forward and up. This is not, however, the whole story, because some /i/s were found which were actually lower than some /ɨ/s, and both were perfectly unglided. Length, as a distinguishing factor, distinctive feature if you like, was also ruled out as length for both nuclei ranged from four to about twelve centiseconds and averaged within one centisecond of each other. Part of the difference, but not all, can be accounted for by facts, such as, all vowels in the speech of my informant were lower and farther back on the chart before voiced stops than they were before voiceless stops. This would be covered by Pike's statements concerning environmental conditioning, but all of the above explana-

tions still leave some embarrassing inconsistencies. The easy solution is that taken by distinctive feature theorists, that is, to assume that there is some unknown distinctive feature which distinguishes these apparently ambigious phones. It is absolutely clear at this point that Jakobson's original statement that all phones of a given phoneme shared some phonological distinctive feature is false. Turkish, for example, and other languages, has a phoneme /r/ which has a fricative allophone, a flap allophone and a liquid allophone. Since distinctive features are far from proven, let us attempt to find out what really happens when a sound comes into the ear of someone who is listening to a conversation.

Returning now to the Markov (this is also spelled Markof, but refers to the same type of mathmatical probability in either case) process. This type of probability is utilized in speech in that some sounds occur in some positions and not in others. A typical example, in initial position, would be the English ng as in the word sing, which clearly does not occur in word or utterance initial position. When the ear receives a sound it must compare that noise with a set of patterns that are stored in the brain and determines which sound (phoneme) this phone represents. In the case of English, this first sound is critical, and probably is in all languages, because once this initial sound has been identified, the statistical probabilities for all of the sounds in second position have been set.

Let us stick, for the moment, with the problem of recognizing that first phoneme. In English this is always, I repeat ALWAYS, a consonant. What about words like "eat?" you say. These begin with a glottal stop. I know that some phoneticians disagree with this, but it certainly is true for my speech and the speech of most of the informants with whom I have worked.

If the patterns are stored in the brain somewhere, then there must be some sort of search

procedure to find a match within the brain. This implies further that the fewer items that need to be searched for, the more quickly a match will be found. Since it seems that consonants, identified by the presence or absence of distinctive features, are more easily matched than are vowels, this should make the identification of the first phone quite fast, because it reduces the number of compare procedures, whatever they are and however fast they are made, which the brain will have to make by the 15 vocalic nuclei in English. That is about 1/3rd of the total phonemic inventory. This should be very helpful, since the vowels are both numerous and difficulty to identify on the basis of sound alone. All of the consonants that begin words in English are easily identified as unique constellations of distinctive features, except the /h/, which is considered to be either the absence of all distinctive features, or the presence of aspiration, both of which are impossible if the initial sound is a glottal stop or some other consonant.

Once the initial consonant has been identified, those phonemes possible in second position are drastically reduced. If, for example, the initial sound is identified as the glottal stop, then the second phoneme absolutely MUST be a vowel. Hence the brain has only to make a maximum of 15 comparisons in order to correctly identify the vowel that occurs in the second position in that utterance. Also, if the first sound is not a glottal stop, but is a voiceless stop, i.e., /p, t, or k/, then the second sound cannot be a stop of any kind, cannot be a fricative, cannot be a nasal or a semivowel. This means that in second position, the brain has only to decide between the vowels or the consonants /r/ and /l/. This reduces the search pattern almost as much as if the initial sound had been found to be a glottal stop.

Once the second sound has been identified, for example, if it is an /l/, then the only sounds that can follow, unless a morpheme has been identi-

fied, no further consonants are possible in that sequence until a vowel occurs, and not all vowels are possible in the next slot. Even more important than this, the phoneme /ey/ as in play is many times more frequent than some of the other phonemes following a /pl/ sequence. If the brain has built into it only a natural sequence of testing for identification of phonemes and some way that searches and attempts to match the incoming sound with a stored pattern which begins with /ey/ and works through the storage of sounds in order of descending probability of occurrence, the speed with which sounds could be identified would be increased enormously. Once the vowel, our example /ey/, had been identified, the brain would need only to attempt three of four matches to correctly identify the next phoneme, and so on.

The concept of the Markov process is that in any sequence of linguistic symbols, either phones, morphs, words, whatever, in any language, given the identification of specific units occurring in a specific order, each similar unit has a statistical probability of occurring in the next slot. That this is absolutely true has been demonstrated over and over again both in pure research, e.g., filtration of certain sounds (Bell Telephone Laboratories) and seeing native speakers fill in the blank slots without being aware that the slot was even blank, and in English teaching programs and materials, e.g., the so-called cloze procedure wherein an element is omitted and students are asked to fill in the blank. The extent to which a foreign student has mastered English is measured by the extent to which he can properly fill the slot in a way that parallels the behavior of native speakers without knowing what the sentence means. Further research (Martin Joos, personal communication) has shown that often, if an improbable sound occurs in a specific slot, it will be misheard as one that is more probable in that position and that situation. The evidence indicating that the Markov process is exceedingly important in spoken

language is overwhelming. There can be no real doubt that this is one of the most important factors in the way verbal symbols are identified in real speech.

That the brain does search for those most used items first, be they phonemes, morphemes or whatever, is indicated by the fact that a tape recording in a foreign language sounds as if the speakers are talking terribly fast when one first begins to listen to the recordings. Later, however, the speech appears to the student to have slowed down enormously. This is a result of the probable fact that the search and compare procedures at first follow the pattern that the student has been using for his native language. This often means that the brain is making comparisons against sounds that are absolutely impossible in a given slot, and this so slows down the identification, that the student cannot keep up even with slow speech in a new language. As the student subconsiously (deep structure?) learns the sequence in which the comparisons are to be made in that language, the tape recorder appears to slow down.

The fact that phonemes, morphemes, and so on do not follow random sequencing in languages everywhere is not an accident of nature. It is an absolute necessity for language systems to work. It is this fact that has forced linguists to include in their descriptions of languages the syllabic patterns of that language. For example, in Japanese the only syllable patterns possible are, V, CV and CVn. There are also at least two places wherein the vowel is often omitted in certain sequential positions, e.g., /suki/ is usually pronounced /ski/, because if the vowel is omitted after an /s/, it has to be a /u/, i.e., in Markov terminology, the /u/ is 100% predictable and as a result the absence does not reduce communication. Other vowels are clearly articulated in that particular type of phonotactic slot. What has been said about Japanese phonemes is e-

qually true for all linguistic elements in all languages, e.g., if an utterance begins with I, thousands of morphemes, and some whole classes of morphemes cannot follow the I without an intervening juncture, which is usually written with a comma.

The situation described above is true for all aspects of all languages and, hence, can be assumed to reflect something in the basic wiring of the human brain. That is, one of the basic elements in the brain which makes human speech possible is a search and compare sequence which flows from left to right (if you will), i.e., in a temporal sequence from the beginning to the end in an ordered mode from the beginning of an utterance to its end, and it is precisely this fact that forces the speakers of all languages to build into their systems limitations on the possible sequences. This is not just an arbitrary, useless piece of objective evidence, it is essential to human speech, because without such restrictions rapid speech would not be comprehensible to another member of the community.

What we have asserted so far, then, and presented some supporting evidence for, is that distinctive features are critical and an important part of all human speech, but that they are not the whole story. We have also asserted that all that is transmitted from a speaker to a listener is a verbal symbol. All research that I have seen shows that people will accept anything that they hear to mean anything that they are told that it should mean. This means that the meanings of any symbol or sequence of symbols is a direct mathmathical function of the life-history of the individual speaker. Hence the only relevant subject for linguists is the way the system of symbols works for the average speaker of the speech community. The concept of "competence" is ludicrous, except as a type of snobbishness which, in my opinion, has no place in modern science. Any one is competent who has lived and utilized a

language, because he can express the things that he needs to express. He may not be able to read (properly, according to an English Prefossor) or compose great literature, but very few people can, especially those scholars who grade everyone as "incompetent" who do not use English (for example) the same way that they do. The only way that competence makes any sense at all is if you are willing to set yourself up as God and then measure everyone else down from yourself. No linguist uses English the same way as a physicist does, nor does a physicist use English the way a chemist does, or a chemist as an English Professor, because each is using the language in a way which best accomplishes what he wants to accomplish. One is as "competent" or as "incompetent" as the other.

Now, how are vowels identified? We have said that they appear to be fuzzy sets, that is, sets which have a core nucleus or central tendency but are clusters which have fuzzy edges. We showed that the fuzzy edges overlap and that the same sound is sometimes recognized by the listener as one phoneme and at another time as another phoneme. We have said that part of the identification process is the Markoff process, part of it a matter of environmental conditioning and part of it the actual phonetic quality of the sound. However, there is still more to it that even this. The listener is not hearing phonemes in a vacuum. He is making choices between possible words. He only has to recognize which of the possibilities the speaker intended. This often means that the actual quality of the sound can range very wide indeed and still be quickly and easily recognized. For example, there are words in English with all of the vowels similar to eat, as in the sequence: eat, it, ate, et, at, Ott, ought, oat, ute, out, but no such words to rhyme with the vowel in book, but, Bert, bout and boy. At least a considerable portion of the fuzzy edges in the sets are a result of the fact that there is no sound

similar to that in but which initiates a word which ends in /t/, comparable with eat. It should surprise no one, then, that the ranges of vocalic nuclei should enlarge and lap over into those areas in which no phoneme occurs in a given environment. An even more serious complication is the fact that there is always a social situation in which a language is being used, and in that particular situation, some words are so unexpected as to be almost impossible for the listener to hear. For example, in my dialect there is a clear difference in the pronunciation of the words latter and ladder. I note this because in some English dialects this difference is not maintained, but it is clear in mine. However, in rapid speech, when the word being used is absolutely clear from context, that difference disappears. Also, in my dialect the words pin and pen are pronounced exactly the same. I cannot make this distinction in normal speech, though over the years I have learned to make it when teaching a linguistic class for demonstration purposes. Fear not, you are no better than I am, if you make the distinction, because if you make that one, you will not be able to tell the difference between the vocalic nuclei in horse and hoarse which in my dialect is the difference between the nuclei in the words boat and bought. If an American speaker of English makes one distinction, he does not make the other, with very very few exceptions. In fact, I have met only one person in years of study who made both. The point is that these differences are usually told by the listener from the social context into which a language is placed. However, going back to the first example, the one wherein I distinguish between latter and ladder, I will make a clear distinction when I know that the meaning is not clear from context. What this means is that the actual sounds I produce in speaking are determined to some extent by the social situation.

If you have been reading carefully, you will

have noticed that some sounds are identified only from the sound of the following phoneme. How is this possible? Very easy. Computers do it all the time. When a sound in a particular range occurs as a representative of two three or even four phonemes at different times, as soon as the brain realizes that this is one of those sounds, it "holds," that is, it does not make an immediate judgement, but withholds judgement until it hears the following sound, which in English will always be a consonant, if the questionable sound is a vowel. As soon as it identifies the following consonant, which will be much easier to do than to identify the vowel, because the consonants are identified on the basis of distinctive features, then it can easily decide which of the three or four possible vowels the speaker intended in that context. Often, however, the brain does not have to wait, because the word meant is already clear from the social or linguistic context.

There has been considerable argument in recent years concerning whether a morpheme is, indeed, a sequence of phonemes or not. I wrote an article in Linguistics, *The Phonemic Composition of English Morphemes*, Volume 20, pp 60-71, 1966, which was criticized by a South African Linguist because "no one considers morphemes to be sequences of phonemes anymore!" Well, so far as I can see the situation, we have no real choice but to consider morphemes as sequences of phonemes, and never have had any real choice. When we deny the existence of phonemes in this sense, we are engaging in a type of philosophical double-talk which is exactly comparable with a chemist who denies the existence of atoms in physical theory. It is, of course. possible to describe any chemical element in terms, such as, electrons, protons, neutrons, positrons, even quarks, and so on, but in so doing, you so complicate the discussions that it would become absolutely unacceptable to any reasonable person, and the elimination of the phoneme from linguistic theory is objectionable on the same grounds. Of course, if you wish, you can talk about a mor-

phemes as constellations of distinctive features, but when you do, you open yourself up to have to face all of the problems described with the very nature of distinctive features and the basic problems of phonemes are simply lowered to that level. It is absolutely clear that the phoneme is a vital theoretical unit when we consider words such as the list below:

/i/	the name for the letter E
/it/	eat
/ist/	east
/its/	eats
/istr/	Easter

and so on. In a stream of noise that comes in, a listener can only identify /i/ after it hears the next sound, as /i/ or /it/, i.e. sequentially. There is no possible way the brain can identify a sound before it gets to the ear, and it is clear that it begins to process the sound as soon as it is heard, so languages must work from left to right, with the basic restriction mentioned above, that is, when a unit has multiple possibilities, the brain simply holds that item (phoneme, morpheme, word, or whatever) until it has sufficient information to uniquely identify it.

It is my contention that one cannot intelligently discuss languages without the concept of the distinctive feature and the phoneme, as well as the concept of the fuzzy set. It is the interplay between crisp sets, identified by distinctive features, fuzzy sets, the social and linguistic context, through a Markov process, as well as the mental set of the listener which gives language its beauty, its flexibility and, indeed, its utility.

Taking English as only one example of languages generally, then, we see from much acoustic research that it is imperative that the initial sound and all syllables under primary stress be correctly identified (cf. both Bell Telephone Laboratory re-

ports and Hasking Laboratory reports among much additional acoustic material). Other phonemes can either occur or not with relatively little observable loss of information. However, it is not as simple as just dropping any phoneme you wish. The reason that phonemes must be considered as a basic, necessary theoretical tool of the linguist is the fact that when a morpheme occurs as a complete utterance, if all of the phonemes are not present, that word may well be misidentified by a listener. If a morpheme occurs with no cultural or social situation to carry information, then the information is carried only through the phonemes. Thus if someone says bit instead of beet, a listener will walk away thinking that the person wants a small piece of something rather than a large red vegetable. In a social context, there may not be a contrasting word with a particular sound in it which makes sense, or one makes considerably less sense than another one, so the listener selects out of his inventory the one which makes the most sense to him under the circumstances. It is this reliance on the social situation which has made many linguists drop the concept of the phoneme, stating that it really is not a significant theoretical tool. However, people who are really interested in communicating with their listeners articulate very carefully in a confusing situation. For example, a person who lives at a race track will articulate the word bed very very carefully if he has occasion to say, "I made a bed at the race track," because he knows that his listener is going to think "bet" if he does not make the sentence clear from his phonemic production. On the other hand, if he is talking to a friend, who also lives at the race track and who also makes his bed every day, he will not feel the need to enunciate so clearly, because his friend will pick up what is being said without the phoneme being clearly articulated. My favorite story is about the Transformational Grammarian who was talking about phonemes in his class one day

and who was asked by a student if it were not true that Transformational Grammarians did not believe in phonemes. His reply was that they really did not, but he didn't see how one could talk about language without them, Q.E.D. The fact is that one cannot talk about language without phonemes. These are the basic phonological units, and the facts given above about their nature is the reason that linguists in the early part of the twentieth century demanded that phonemic contrast be proven by minimal pairs, preferably single morphemes, preferable utilized as a complete utterance, because only in that situation could one be reasonably sure that the listener could get the information only from the phoneme. However, until very recently we did not know the true nature of phonemes and fuzzy sets.

As stated earlier, the initial phoneme is identified by distinctive features, and at least in English, is always a crisp set (i.e., a consonant). This sets the Markov probability for all of the phonemes in the second position, which greatly facilitates the search and identification in the brain. If the second item is a consonant, this too then is a crisp set, except possibly for /l/ and /r/, but once this second phoneme is identified, this greatly reduces the possibilities in the third position. Finally, one will eventually come to a vowel, which is not a crisp set, but which is a very fuzzy one. This fuzzy set is not identifed on the basis of distinctive features, as were the consonants, but a sound is isolated, then the listener must choose between various possibilities on the basis of a combination of features, for example, shared life experiences, the fact that no word exists in his vocabulary with a particular phoneme in it, and so on. Finally, when a morpheme is completely identified, all of the Markov probabilities are reset, in terms of what can possibly occur in the next syntactic slot. What this means is that after each morpheme, the probabilities are not automatically reset as if one were starting a new

utterance.

What does what has been said above tell us about the nature of the basic wiring of the human brain? First, it tells us that the information on all languages studied throughout the world so far show severe limitations on the sequencing of phonemes. Are there any limits on specific sounds? NO! I personally have worked with languages in which no sequences of consonants occur and in languages in which sequences of seven consonants occur within a single moprheme. If one strings morphemes together, one can get considerably more than seven. However, perhaps there is something in the brain that sets seven or eight as the upper limit on such sequences within a morpheme, because one can string morphemes together, and perhaps the brain does not feel comfortable handling so many consonants. This is by no means certain however. Further, some languages restrict vowels to single vowels surrounded by consonants, except across morpheme boundaries. English tolerates very very few vocalic sequences. Other languages, such as Japanese or the languages of Polynesia have long strings of vowels within morphemes. THE ONE THING THAT APPEARS CLEARLY TO BE A LINGUISTIC UNIVERSAL IS THE FACT THAT ONE CANNOT SIMPLY HAVE A RANDOM SEQUENCE OF SOUNDS. The sequencing must be ordered. I believe that this is the result of something within the human brain and that it is found in every language because of this fact. This is very different from saying that all languages have nouns, because the finding of nouns is most likely a product of the methodology of language description, and I do not believe that the possible sequences of phonemes is a product of the method of description. Almost every possible type of sequence occurs, for example, Wayam, an American Indian Language has many sequences of rearticulated stops, both the same stop and different stops. English does not tolerate the sequencing of stops except across morpheme boundaries, and even across morpheme

boundaries, one of the two stops is likely to be dropped in normal speech, so strong is the native speakers aversion to such stop sequences. Words that begin with a /tt/ cluster simply cannot occur in English, but they do in other languages. Such information on many languages demonstrate that the only general thing about sequencing is the fact that it occurs. This must be so because without it there could be no Markov process, and without a Markov process, there could be no rapid speech. Random searching of the storage compartments in the brain would simply be too slow, and having the sequences prescribed facilitates controlled, limited access to this storage and makes rapid human speech possible.

What we have stated in this chapter is that all languages have phonemes, in fact, all languages have two types of phoenems, consonants, which are crisp sets, and vowels, which are fuzzy sets. The crisp sets are identified in speech through the recognition of distinctive features, as a rule, but a redundant consonant, or one which is not needed because of context, may well be omitted in normal speech. This fact has led many linguists astray in their attempt to understand how languages actually work. Vowels are fuzzy sets and their identification is much more complicated than is that of consonants. That is precisely the reason that they have been the topic of so much controversy over the years. These fuzzy sets are identified through a variety of things including context, Markov probability, shared life history, and so on. Further, we say that all one needs to postulate as a part of the human brain's natural wiring (deep structure, if you like) is that it is set up in such a way that it can store patterns for the identification of sound and that these are always set up so that the search can procede in an orderly manner and not at randon with the search beginning at the most commonly utilized phoneme in that syntactic position. Further,

all languages have a smaller proportion of vowels than consonants, because of this difficulty in identification of the fuzzy sets. The fluidity of language, and indeed its beauty and flexibility in speech is a direct result of the interplay between the crisp sets, which give it a rigidity and stability, and the fuzzy sets, which give it a flexibility. Too much ridigity would take away much of the artistic usages we love so much and too much flexibility would take away its utility, which is so necessary for the efficient functioning of a human society.

CHAPTER III

In the last chapter we saw the way that natural languages work on the phonological level of analysis in a real speech situation. Now it is time to begin to concern ourselves with the meaningful symbols of language. The bottom of the line so far as meaningful units in language are concerned is the morpheme. The morpheme is the shortest sequence of phonemes which carries a meaning. This means a sequence of phonemes which cannot be subdivided without losing its meaning. There is nothing new in this definition. It is the same as linguists have been using since the latter part of the 19th century.

It is my contention that virtually no language so far described, except a few American Indian languages and some of the Indo-European languages closely related to Latin, have been described even remotely as they should be. The problem stems from a fact relating to linguistic typology. Sapir's typology is a good starting point. His classification of languages into isolating, aglutinative, synthetic and inflective will serve to illustrate the problem. Going back to the days of the Greeks, they studied basically their own language, and the theoretical tools which they utilized or developed were suited to the analysis of highly inflective languages, because Greek, Latin and in fact virtually all Indo-European languages are highly inflective. Grammar, as a result of the hitory of linguistics and grammatical analysis, is often defined as the process of word formation, and many linguists speak of morphemes only as if they were

suffixes, prefixes or infixes. In fact, for those who have seen the series of films on *Languages and Linguistics*, which were produced by the Institute for Human Communication, State University of New York, Buffalo, will remember that Henry Lee Smith, the narrator, states explicitly that in all languages there are two fundamentally different types of morphemes, stems vs affixes. The importance of this statement is not that it was made by Dr. Smith but that it reflects a longstanding conviction among many linguists, and accurately reflects the status of linguistic thought, even as late as the mid 1970's.

Do all languages have two fundamentally different types of morphemes? Of course they do! However, they are not stems vs affixes, as is often stated in the linguistic literature. This myth comes to us directly from the Greeks and has not been modified significantly over the past two thousand and five hundred years. The concept is a result of the basic fact that Greek was a highly inflective language, and naturally the thing that the Greek grammarians saw was that in their language there was this fundamental dichotomy. When the concepts for grammatical analysis were transmitted to Rome and the Latin grammarians, this appeared to be a valid dichotomy, because Latin too was a highly inflective language, hence much that is grammatical was handled through affixes. During the middle ages and even into the nineteenth century, linguists still saw this as a basic dichotomy, because they handled only the European branch of Indo-European languages and other highly inflective languages. Every time a linguist found himself faced with a non-inflective language, he had what appeared to be insoluble problems. No attempt was made to work on this problem until C. C. Fries began to look more or less objectively at the structure of English in the 1930's (cf. Charles Carpenter Fries, *The Structure of English; An Introduction to the Construction of English Sentences*, Harcourt, Brace and Co., 1952

for further details on this). Fries was the first human being on the face of this earth to tackle the problem, at least in print. He asked the questions about English grammar that should have been asked centuries before but were not, because of the dynamic control that the language and thought patterns of Indo-European scholars held over the speakers of Indo-European languages. He asked, for the first time, what are the parts of speech in this language, instead of saying, what are the nouns, what are the pronouns, and so on? The point is that for over a thousand years, the structure of English was predetermined by the questions that were asked about the language instead of being determined by the actual way the structure of the language functioned.

Fries was still too much of a traditionalist to ask all of the correct questions, because he, like most of us, had been trained, that is, brainwashed, by years and years of education in English departments, where people do not ever ask questions about the nature of English. The English Professors are not really interested in English. They are interested in literature. That is why they go into English. It is this lack of interest, not their bullheadedness or any other unreasonable thing that has led to their not asking the correct questions. Unfortunately, Fries did not come up with a workable solution for the problem of analyzing isolating languages. Note here that the problem basically is that isolating languages work in a fundamentally different manner from inflective languages, but since there had been no long term study over history of any non-inflective languages, linguists did not even understand that there was a difference. Fries postulated a third type of morpheme, the function word. However, this, with the passage of time, was generally found not to be satisfactory for the handling of languages generally, but it was a step in the right direction. It was the foot in the door, so to speak, which would eventually open that door to insights into the true

nature of linguistic analysis.

Quite simply, the problem was that for centuries European linguists and grammarians had handled no languages that were not inflective to any great depth, and as a result had no idea of how these languages worked. In 1951 I was a part of a seminar at Indiana University, led by Carl F. Voegelin, in which we tackled the problem of working on a type of analysis which might work for all languages instead of just those that were highly inflective. We did not, at the time, come to any great conclusions, except for Voegelin's suggestion that the real dichotomy was between what he called Major and Minor morphemes. He did not, to my knowledge, relate this specifically to isolating languages, but he recognized the basic fact that function words were not the solution to our problem, because there was, indeed, really only two types of morphemes in all languages, but that in many languages these were not all stems or affixes. That they were stems and affixes in European languages resulted from the fact that European languages generally descended from a common parent, which had been highly inflective and the apparent dichotomy was a product of this historical tradition and not a fact of language design or function.

What happens now is that we are again back to a dichotomy, but at this point in time we can handle all languages whether they are isolating, agglutinative, inflective or synthetic with these two categories whereas with the old system we could adequately handle only inflective and agglutinative languages, that is, languages with a multiplicity of affixes. This was true, because morpheme classification was determined by setting up classes of words that could follow the same sets of inflections and sets of affixes which could be affixed in exactly the same manner to a large number of stems. If you follow this old concept of grammar, that is, that grammar is the process of word formation, you come to the completely untenable conclusion that languages such as English

and Chinese, that is, languages which are isolating, have no grammar. In fact, in some French schools today I have been told by French students, they are taught that English has no grammar. This is a direct result of studying languages through the wrong end of a telescope; a telescope which was built by studying languages all or most of which are a single type.

What does all of this mean in terms of the theory of how languages work? It means that we must analyze the languages in terms of something other than stems vs affixes, and that one of our first problems is to break the collection of morphemes in a language into two great groups, those which are major morphemes and those that are minor morphemes. If you are dealing with a highly inflective language such as Latin or any of the Romance langauges, the old method will adequately handle well over 90% of the analysis, but still leaves some unfortunate problems. With highly agglutinative languages, such as, Turkish, this method will also work, because there are so many affixes that it is possible to handle a great deal of the language structure following the traditional procedures. However, there are still a great many things even in Turkish which are isolating in nature.

The last statement in the paragraph above leads us directly into a problem not faced until about 1950. That is, Greenberg's recognition of the fact that probably all languages have some morphemes which are isolating, some agglutinating and some which are inflective (cf. Joseph H. Greenberg, *A Quantitative Approach to the Morphological Typology of Language*, in Robert Spencer (Ed.), *Methods and Perspectives in Anthropology*, University of Minnesota Press, 1954, pp. 192-220 or IJAL:26: 178-194, 1960 for details). The importance of the recognition of this hitherto ignored fact about languages is that the traditional method worked to exactly the extent that a language was inflective, and since languages were not either inflective or

non-inflective, but inflective to certain degrees, any grammatical analysis worked out on this basis would be accurate to the extent that the languages were inflective.

How could we then adjust the method of analysis utilized for centuries so that it would be able to handle languages with few or no affixes? That was the question for linguistics in the 1960's. Think what affixation is for a moment. It is a particular way of sequencing. Linguists have always felt that the relationship between a stem and its affixes was a closer relationship than one between an article and the following noun, but it really is not. It is the same. The difference is only in the nature of the language being studied. If it is an isolating language, then we have to use sequences such as "it is a _____," to make our analysis. Does this look familiar? Of course it does. This is a syntactic slot, and what we are really looking for when we attempt to classify morphemes are groups of morphemes which will fill the same set of syntactic slots. What one was doing following the traditional method of handling morphological analysis was to classify a group of words, called nouns, into one classification because they could all be the subjects of sentences, the objects of verbs, and so on. It was basically never that they were the names for things. After the fact of analysis in Latin and Greek, it was determined that many, if not most, of these words were, indeed, the names for things. However, it does not logically follow that these same types of things have to be so in other languages. The same thing is true for verbs or for any other part of speech.

When one begins an analysis, he has to start somewhere, and to a certain extent this is on the basis of his intuition. One of the great debates over the past twenty years has been the extent to which intuition needs to be utilized in analysis and the extent to which some sort of rigid methodology should be utilized. Emmon Bach says, "... in

linguistics... what we must account for includes what is known as the native speaker's intuitions about what he says and hears..." (*An Introduction to Transformational Grammars*, Holt, Rinehart and Winston, 1964, p 3): Now, as a general statement, probably no linguist would disagree with that, at least not in principle. The problem comes when we attempt to interpret exactly what is meant. So long as we are talking in terms of an unsophisticated native speaker, one who has not been given a long period of education in formal grammar, no one would argue the point. The problem comes when someone like Chomsky, or, indeed, Bach himself, offers intuition as the principal basis on which an analysis is made, and this intuition is strongly colored by traditional grammar which they have learned in school.

To illustrate what is meant, if we take a good look at Japanese grammar, we find three pronouns, *watashi*, *anata*, and *kare*, which mean I, you and not I or you respectively. All of the grammar books that I have seen, classify these as pronouns. Yet on what basis are these called pronouns? The basis is that they translate into European languages as what are pronouns in those languages, and such a classification has nothing to do with Japanese. How does this relate to intuition? If you ask any educated Japanese what these forms are, he will tell you that these are pronouns. Does this reflect a native speaker's basic intuition about the nature of his language? Absolutely not! The original grammatical analysis of Japanese was done by Europeans on the basis of European logic and grammar. From any reasonable point of view, these forms, *watashi*, *anata*, *kare*, and the social variations of these, such as, *boku*, *kimi*, are very clearly nouns in Japanese. What I mean by this is that these forms do exactly what all nouns do, and they do all of these things in exactly the same ways that nouns do them. There is absolutely no logic of any sort in calling these pronouns, except that they translate into English and other

European languages as pronouns. The Japanese native speaker will call them pronouns, not because he feels anything about the language, but because he has been taught in school that they are pronouns, and that analysis and description comes from the same grammatical tradition which makes English and, indeed, the grammar of most languages around the world, unintelligible.

What do we mean when we say that these are nouns? We mean that we do not have a singular and plural form, as we do in English, such as I as opposed to we, rather we pluarlize the form watashi by adding the noun pluralizer to get watashitachi. If we pluralize friend, tomodachi, we get tomodachitachi. Note that the pluralizer is exactly the same for the so-called pronouns as it is for the noun. Further, if we possess things in Japanese, we say onna no te, the woman's hand, and we also say, watashi no te, my hand. These pronouns can also be possessed, as in anata no watashi (literally your me) meaning I am yours. There is no point in pursuing this point of grammar. The fact is that anything you can do with a noun, you can do with the so-called pronouns, and logically this absolutely implies that these forms are nouns. One of the salient differences between English, which is completely obscured in traditional grammatical terminology, and Japanese is that English has four sets of pronouns, which are minor morphemes, and, indeed, are quite a different kind of grammatical unit from nouns whereas in Japanese there is absolutely no class pronoun which is different from that of noun.

What we are trying to do, then, is to develop a method of describing and analyzing languages which will point up salient differences and similarities so that eventually we can, by comparing the descriptions of languages, find those things which are really linguistic universals. However, if we look at present descriptions, we might well conclude that all languages have pronouns, and

it is perfectly clear that such a conclusion comes from the way we describe the languages and actually obscures the fundamental difference between English and Japanese, as well as many other languages.

We stated earlier that there are two fundamentally different types of morphemes in all languages and these are labelled major and minor morphemes, not stems vs affixes. Essentially the difference between the two types is that the former compares with what have been traditionally called lexical morphemes or what Fries called content words. These latter forms, that is, the major morphemes occur in very large classes, usually hundreds at least, are very fuzzy sets, that is, the words included within that set (say nouns) can change drastically, even in a very short period of time and finally, they label the things that a speaker from a particular culture wants to talk about. The major morphemes differ from minor morphemes in that minor morphemes come in very small classes, usually less than a dozen, but occasionally slightly above that, have grammatical meanings and remain very very stable over very long periods of time, say hundreds of years. Also, at least a very large number of minor morphemic sets are traditional crisp sets which can be described in terms of what have been called semantic components, that is, meaning-distinctive-features if you will. One further connection between crisp and fuzzy sets, native speakers may disagree as to whether a certain morpheme can be used as a noun or not, but they will not disagree as to whether a certain pronoun is different or not. So we see fuzziness and crispness exhibited in two ways for major and minor morphemes. Hence it would appear that the concept of a fuzzy set theory for at least part of language description makes very good sense.

Relating all of this to types of languages and the problems revolving about analysis, we see that traditionally classes have been set up on the basis that certain suffixes went with certain stems.

It now necessary to alter the basis of our analysis to that of various sequences, at least for isolating languages. This then should give us a picture of isolating languages which is parallel with that yielded by traditional linguistic analyses of inflective languages.

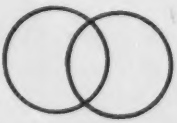
Taking English again, as our example for an isolating language, we see that we can set up the frames:

I am _____-ing and It is a _____.

To isolate what we have been accustomed to calling nouns and verbs. If we begin to substitute words that we know into whichever frame they will easily fit, we will get a list which we can call nouns and a list we can call verbs. Then the question arises, does such a classification yield viable classes of morphemes in English. This is not as easily determined as one might at first imagine.

I used the above method of analyzing in my paper which appeared in Linguistics, i.e., *The Morphemes of English: Major Morpheme Stem Classes*, Linguistics, Volume 36, 1967, pp 29-51, and found that when I set up the class noun and the class verb and made a list of the morphemes in each class, just over 93% of the two lists were the same. If we set up two classes which have an overlap of 93%, it is questionable that this set of classes is valid. Note the three situations below wherein classes are represented by circles.

Two Separate Classes



Three Separate Classes

Only A Single Class

Relating the sets of circles on the opposite page to the classification of English morphemes, we find that nouns vs verbs form a single class. Linguistic analysis is the collecting of morphemes into groups which perform similar or identical syntactic functions, and this is a problem in classification, nothing more. If we have the situation seen in the two circles that do not intersect, we have two clearly distinct classes. This is typical of the situation in Turkish, wherein words are c'early nouns or verbs, and one cannot utilize the langauge correctly nor correctly identify the meaning of a sentence, unless we know whether the words utilized in the utterance are nouns or verbs. In the second pair of circles, that is, the two that intersect, we see a language in which we have a set of morphemes which we might call verbs, one which we might call noun-verbs and one which we might call nouns. This latter picture is the one that I was brought up on in English grammar classes. The idea was that in English there was a large class of nouns, a large class of verbs, and a small class of words that were sometimes used as one or the other. However, the real situation is not this at all. The real situation is that there is a 93% overlap of the so-called classes of nouns and verbs. That means that the picture showing two circles which overlap almost 100% is the correct one, and this clearly shows two classes which are really only a single class. That is, the real situation in English is that the langùage does not make a distinction between nouns and verbs. The proof of this can easily be found in events such as those I will describe below.

When I collected the original data, I listed a small number of words as pure nouns, one of these was the word window. A few years after the original collection of data, I was explaining this to a class and one of my students raised his hand. "I window houses every day," he said. When asked what he meant, he stated that he was a carpenter's

helper, and it meant that they were putting in the windows of the houses that they were building. This passed, and I thought, so I made one mistake. Then a few months later I was driving home in my car from work and I turned on the radio. Out comes the words, I'll be Christmassing with you. I had always thought that Christmas was as clear a pure noun as one might ever find, but there was clear proof that the people who speak English as a native tongue did not consider it to be so. Not too long after that, my wife said to me one night, would you please 'lid' the jar, and so, finally, after several years of trying desperately to save the traditional grammar on which I had been raised and in which I believed so strongly, I was forced to admit the basic fact that the native speakers intuition tells him that there is no distinction between nouns and verbs in English grammar. These and other events, based on the behavior of English speaking adults in real speech situations, finally led me to the conclusion that the most significant class of major morphemes in the English language is the class noun-verb. The explanation for the few morphemes which cannot be utilized as both is simply that they are the exceptions that one seems always to find. There is absolutely no reason that every single noun and verb should be utilized as both. The ones that we do not currently use as both are simply those for which no nominal or verbal meaning has as yet been found. For example, ocean seems to be a pure noun, and I cannot imagine what one might mean by 'oceaning' anything, but I am very sure that when someone does use it, and I hear it or see it in a linguistic context, I will know exactly what the speaker or writer means, because we as native speakers do, indeed, not make a distinction between these two classes of words. If we did, we would have a very unwieldy grammar, because we would have many thousands of words such as walk, talk, lip, mouth, head, and so on which are utilized as both nouns and as verbs with essentially the same meaning, and even simple rationality tells us that this is not

a reasonable way to classify morphemes in a linguistic system.

Having determined that nouns vs verbs is not a valid classification for English morphemes in terms of the way native speakers use the language, we are then faced with the question, are there any other morpheme classes beside the one noun-verb or verb-noun, whichever appelation you prefer? Yes there is, because we can put a large number of morphemes into the following frames into which members of the class noun-verb do not fit.

This is _____-er. This is _____-est.

Into these frames will go such words as: tall, big, small, tiny, and so on. There is some overlap between this class and the first class established, and this brings us to some of the problems of analysis which are rarely touched upon by most linguists. That problem has to do with statistics.

When I say statistics in linguistics, I do not mean complicated statistical formula, but simple counts. Glancing back at the three sets of circles presented to show why English had only one class, i.e., noun-verbs, rather than a class noun and a class verb, we see that it is only through examining the counts of morphemes on various lists that tells us whether we are dealing with a single class or more than one class. This is true because of the fuzzy nature of some linguistic classes. A great many linguistic classes are fuzzy. Certainly all major morpheme classes will be fuzzy because of the fact that new words are coming into these classes and old words are being lost from these classes almost constantly. Hence, only by counting can we tell that setting up a class of morphemes called nouns, as opposed to verbs, makes no sense. Unfortuantely, I have not had the time to do the same thing for the class which fits into the above frames as I did for noun-verbs, so I do not know for sure that the forms which occur as both noun-verb and modifier (as we shall call this

class) are valid sub-classes or just the fuzzy edges of the class, i.e., elements in the fuzzy edge can be called exceptions if they are not numerous. However, as the length of the list increases, it will become more and more clear that the people who speak the language must distinguish this list into a separate class which would by definition then be a sub-class. The label modifier, to replace the two classes in English which have been traditionally called adjectives and adverbs, is not original with me. It has appeared in several other publications, e.g., *The Outline of English Structure*, George L. Trager and Henry Lee Smith, Jr., Studies in Linguistics, Occasional Papers, No. 13, Battenburg Press, 1951. It is perfectly clear from hundreds of examples that there is a class of modifiers in English, but that these are not divisible into adjectives and adverbs.

Long in the past it is probable that there was a division into adjectives and adverbs, but at the present time, even the educated speakers say such things as,

The fast man ran fast. (as opposed to quickly)

Clearly, modifiers would not be expected to be divided into adjectives which modified nouns and adverbs which modified verbs if there was no classification into nouns and verbs, though this is not necessarily the case, because adverbs could be those morphemes which modify predicates and adjectives could modify everything else. However, in English, as it is utilized, even by educated speakers, there is no class adverb. A modifer modifies nouns if it precedes the nouns and predicates if it follows them. Meaning is understood completely from syntactic position and not from the class of the morpheme. We see this in the example above.

What function, then, does the suffix -ly play in English structure? Clearly it is an inflective morpheme which can be utilized in situations when the modification of a given form could be ambi-

guous, such as in the example, the slowly sinking ship, which clearly is different from, the slow sinking ship. If the -ly is not present, it is not clear whether the ship is slow or the sinking is slow. However, with the -ly affixed to slow, that tells us very clearly that it is the sinking that is slow, that is, the modifier actually modifies the following modifier, and not the noun which is the head of the construction. This does not mean that we have to postulate a class of major morphemes which would be called adverbs at all, no more than the inflection of a noun for genitive case would mean that we had to have two classes of nouns. I know that some readers will immediately scream that the -ly is not inflective but derivative, but they are simply wrong.

The basic difference between inflection and derivation is that a form which has a derivative suffix or prefix added to it is basically different from the stem without such affixation, whereas, the stem with an inflection affixed remains essentially the same, but the inflection tells us something about the morpheme, that is, it has a meaning. The kinds of differences that derivative morphemes make are often that a noun is derived out of a verb, in languges which have nouns and verbs, e.g., Turkish, the word boş is a noun meaning empty, but the word boşal(mak) is a verb and it means divorce. Hence we see that derivative affixes often alter both the classification of a morpheme and its meaning. Minimally, derivation must alter one or the other, otherwise it is inflection, which is the addition of a constant element of meaning, such as, plurality, to the meaning of the stem morpheme. Clearly, -ly does not alter the meaning of the stem to which it is affixed, nor does it alter the classification, because one can do exactly the same things with a stem morpheme with the -ly affixed to it that he could do without that suffix. This is an inflective morpheme which says, "I do not modify the following noun, as you might think. I modify

something else in the sequence." That something else usually is a following modifier, or in certain situations, it may be modifying the predicate.

Now, imagine for a few minutes, if we do not know what the major morpheme classes are, i.e., if there is, indeed, no class adverb as such nor any class verb as such, what sort of sense does it make to try to teach people that an adverb modifies a verb, an adjective or another adverb. We cannot possibly make use of this sort of information. Such definitions make sense only if the language has in it a viable class verb, which English does not, a class adjective, which English does not (at least not in contrast with adverb), and a class adverb, which English does not. How, in the name of God is this sort of analysis and information going to be of any use to anybody on this green earth? It clearly is not only not going to be of use, it is going to be confusing and problem causing for learners of this language whether they are native speakers or foreigners who need to use English to communicate with someone. Such statements are the direct result of "intuitive" analysis of English, because when we rely on our intuition, we simply reflect what we have been taught in our English classes. If we have been taught the same thing, our intuitions will be similar. If we have been taught different things, our intuitions will be different. It is absolutely essential that we develop methods of analysis, that is, discovery procedures, by which we can "discover" how each language really works. Statements to the effect that the structuralists have had no success with this are simply not accurate. What the authors of such statements really mean is that the discovery procedures do not discover what we have intuitively felt for many years, therefore they are incorrect. This is not the case. What is incorrect is the intuition of people who hold to the beliefs that English has nouns, verbs, adjectives, and adverbs as classes of major or lexical morphemes. I make no claims

at this point to have discovered all of the major morpheme classes in English, but only to have demonstrated conclusively how foolish intuitive analysis is and how absolutely wrong the traditional analysis is. What I hope I can do in the future is to get young scholars interested in delving into the structure of English so that we can find out how it does, indeed, work. From such studies, perhaps the next generation or two will be able to have descriptions of English which will help them to learn or use the language rather than to confuse them.

To summarize what we have demonstrated about major morpheme classes; English does not distinguish between nouns and verbs. What this means is that knowing that a word is a noun or a verb in no way helps a student utilize the English language in normal conversation or in writing a composition. On the contrary, if languages do distinguish between nouns and verbs, as Turkish does, then that bit of knowledge is not only helpful, it is absolutely essential to proper usage of that language. On the other hand, in Turkish there is no distinction between modifiers or adjectives and nouns. All words which would translate into English as adjectives are nouns and function in exactly the same manner as nouns. Teaching someone about adjectives in Turkish is confusing and harmful to their using the language; just as teaching English speakers about nouns and verbs does not help. The confused state of English grammar stems from the fact that instead of correcting a very bad analysis into fake classes, because those forms would have translated into Latin as those classes, we have tended to say that every word is useful as every possible class, and this is not the way languages work. Each langauge has a set of real classes that make sense in terms of the structure of that language. When one discovers those classes and describes them accurately, this information will make it much easier to teach and to learn that language. As we teach English,

the student is left to make his own analysis, because he finds out almost immediately that the information fed him in the classroom about the language is no help. He then makes his own analysis, exactly as he does in learning his native language when he is a child. Eventually, he unlearns all of the bad grammar that has been taught him and learns to use the language correctly, but this is a long slow process wherein he is taught many things that are not true. He has then to unlearn those things, and figure out the correct ones for himself. What we hope for in the future is that an accurate description of English will make it possible for us to teach directly, from the very first, that grammar that the student must use when he speaks or writes in English. This is the reason we must know "the nature of natural languages."

What we have said so far, then, is that over the centuries since Plato we have basically performed linguistic analysis on the basis of the sets of suffixes or prefixes that stems could take. This is fine for languages like Turkish, and there is little wrong with the analysis of that language. However, even this has been tainted by Latin grammar, because the form which is labelled the "accusative" has nothing to do with the direct object of verbs. This form is equivalent to the definite article in English, and it usually occurs on a direct object simply because one does not hit things in general, he hits a specific one. For this reason and for this reason alone the definite article is usually affixed to the direct object. However, it is perfectly clear from the examples below that this is not an accusative ending, though one can call it accusative if he wishes, so long as he realizes that it is a definite article.

Elma aldım.	I took an apple.
Elması aldım.	I took the apple.
Elmalar aldım.	I took some apples.
Elmaları aldım.	I took the apples.

In each case the word for apples is accusative, but

each sentence has a different meaning, and it is absolutely clear that the so-called accusative endings /ı/ and /sı/ mean a definite item or group of items, not the accusative case.

This, then, brings us to the problem with non-inflective languages, such as English. How do we analyze these into classes, obviously not on the basis of inflections. We stated earlier that this must be done on the basis of slot-filler capabilities, and we showed what happened to the traditional classification into nouns, verbs, adjectives and adverbs when we applied this type of methodology. If we now relegate all "uninflectable" stems to the class particle, we find that English has very little grammar, and the same would be true for any isolating language. The next question that must be answered, then, is, are there any more classes of major morphemes? The answer to that is that it is uncertain, but clearly most of what have been thrown into the class particle in the past now become minor morphemes, and they can only be understood by looking at the paradigms into which they fall. However, if there are other classes of major morphemes, they must be discovered through a methodology similar to that pictured above and not through the misleading intuition of any well-educated English speaker whose intuitions have been warped through years of brain-washing in traditional grammatical classifications.

The nature of major morpheme classes then is that of a large group, usually several hundreds, all of which perform the same syntactic functions. We have also seen that a classification such as "nominal-function" is not valid for English. We lump together the prepositional objects, with verbal objects with subjects of sentences, and we call all of these nominal functions, when we are really talking about three totally different functions. We must define a major morpheme class in terms of the functions that it can perform. We must first find a large group of morphemes all of

which perform all of the same functions. These functions must be something as clear as "prepositional objects," "verbal objects," "sentence subject," and so on. Then we define our class by saying that the class noun-verb performs in English, the subject, verbal object, prepositional object, possessor and possessed syntactic functions, for example. If we do not do this, we find ourselves right back with defining anything that will perform a nominal function as a noun, and we know no more about how English really works than we did before. What a person needs to know when he is learning a new language or interpreting something in an old one is, is there a real class in this language? What is the class? What can I do with it? In Turkish, if you do not know that a given word is a noun, you will miss the meaning of a sentence, especially in so-called sentence-fragment utterances, because the word may well occur without any inflection attached to it, and if you speak English, you might think that the referent was simply a verb being used as a nominal, whereas the word in question can only be utilized as a verb. Hence the correct interpretation of the meaning of an utterance hinges on your knowing that certain forms (stems) are nouns and certain others are verbs. In English this is not the case. In English, we have the same situation with adjectives and nouns, if you are a Turkish speaker. The point is that each language has a unique set of major morpheme classes, and these must be understood by the user. When we describe all languages as being the same, we obscure the very essence of each langauge. Further, one of the basic characteristics of major morpheme classes is that they are fuzzy sets. By this we mean that the composition of the class changes from time to time and is a little different even at a given time for different speakers of the language.

Now we need to move on in our discussion to a description of the nature of minor morphemes. In terms of meaning these have traditionally been con-

sidered to have a grammatical meaning. This is supposed to mean that the morphemes show some sort of relationship between words in the sentence or some word in the sentence and something in the real world rather than refer to or symbolize something. Examples of the differences become clear when we look at a word like house a major morpheme and I a pronoun. The former symbolizes a large collection of things in the real world and the latter relates the action of the predicate to the speaker; the speaker being then the subject of that utterance.

We now need to find a way to classify these into classes similar to those for the major morphemes, because it is clear from even a cursory glance at English that these cannot be set up into classes on the basis of the stems to which they are affixed, since they are, indeed, not affixed at all. This problem arises with all isolating features in any language, and the more isolating a language is the more of the minor morphemes will either be left unanalyzed or misanalyzed if we do not have a proper procedure for identifying these morphemes and collecting them into proper classes. Here again, we rely on syntactic function, since this is basically what we were doing for inflective languages when we discussed affixation. We are all familiar with the series of English pronouns listed below, and I will cite the Turkish equivalents opposite them to show the difference between isolating languages and highly inflective or agglutinative languages.

I ran.	Koştu-m.
You ran.	Koştu-n.
He ran.	Koştu.
She ran.	Koştu.
It ran.	Koştu.
We ran.	Koştu-k.
They ran.	Koştu. or Koştular.
	Koştu-nuz. (You formal)

Note that English yields the set on the left below, and Turkish yields the set on the right below.

I	-m
You	-n
He	-# (zero morpheme)
She	-k
It	-# or -lar
We	-nuz (formal, not pl.)
They	

Why do we say that there is such a difference between these two languages? The difference is profound and very important. The English forms can be utilized as complete utterances in response to questions, such as, "who is there?" The response would in English be "I," but the form -m cannot be utilized alone as a complete utterance. There is a free form, in Turkish, which is not a pronoun but a noun, that is, ben, which must be utilized when there is no verb in the sentence to which the suffix can be affixed. The form ben is a noun because it is inflected for nominative, genitive, and so on, but the pronouns are not. Note also that neither English nor Turkish distinguishes a singular from a plural form for the second person, despite the fact that most grammar books in both languages show them as having plural second persons. The reason that it is totally unreasonable to consider you as a singluar and you as a plural in English is that no one can possibly tell from the structure of the form whether one is talking about a group of people or a single person. Hence, if we are, indeed, interested in understanding how the human mind works when it is dealing with language, we must realize that the difference between singular second person and plural second person in English comes from the social situation or context and not from anything that appears in the symbolic system which is to transmit the messages. In Turkish, there are two different forms, but the second one is utilized

only when one is talking to someone who is his equal or superior in the social order or a group which includes someone above him in the social order. The -n form is utilized for any group in which you are certain that all members are beneath you, but -nuz is utilized for a single person or a group which might possible have an equal in it, unless you are very familiar with the members of the group. In other words, the distinction between the two forms has nothing to do with singular vs plural but social status. In English no such distinction is made. Japanese, on the other hand has several different forms, seven to be exact, for each so-called pronoun, each of which is utilized in different social situations to destinguish between people of different ranks vis a vis the speaker. However, none of the Japanese forms are true pronouns, all are nouns, as described earlier, for which the plural forms are, indeed, possible and are formed exactly as any other noun is pluralized. Hence, Japanese has no minor morphemes which could be called pronouns. Turkish on the other hand has at least one (actually more, but I do not wish here to describe the structure of Turkish) paradigm of minor morphemes which are true pronouns, that is, the suffixes described above, and one group of so-called pronouns which are nouns just as has Japanese. In English, all of our pronouns are minor morphemes and are, indeed, true pronouns. That is, there are no so-called pronouns in English which are actually nouns.

Note in the paragraph above, we see again that the traditional way of handling grammatical descriptions obscures fundamental differences between English, Japanese and Turkish both on the major morpheme side and the minor morpheme side. Then what happens is that linguists scour the world and see that all languages have "pronouns," when they most certainly do not, or if you say that they do, you are saying that totally different things are the same in different languages. It is this type

of misanalysis which leads to misinterpretation of linguistic structures which ultimately leads many linguists to postulate a deep structure which is totally unrelated to what one must certainly find if he could really get inside the head of a speaker of any given language, or a human head anywhere for that matter. When we really understand the basic nature of hundreds of languages, then, and only then, can we have an accurate picture of the "deep structure" or the basic wiring inside human brains.

Why did I make such a point in the paragraph above of the fact that certain elements in Japanese are not pronouns, and that certain Turkish forms are not pronouns, while English has sets of pronouns which are really pronouns? The problem is that it is irrational to set up two classes, on the basis of meaning, when all members of both sets do the same things syntactically. If all morphemes of a given group perform syntactically in exactly the same way, it follows logically that the speaker considers them to be the same thing. Hence we are forcing an arbitrary, and actually useless, distinction on learners of the language or speakers of the language. A person does not need to learn a bunch of nonsense about pronouns when he is learning Japanese, because the information will not help him use the language. When learning Turkish, on the other hand, if a person is to utilize the language correctly, he simply must learn at some point that there are things called pronouns which function as pronouns should, as well as a group which function as nouns. In English, finally, we need only know that there are pronouns, and that they function in certain ways. This information, drawn from a careful analysis, most likely is a very accurate reflection of the way the forms are handled by a native speaker when he learns the language from the ground up as a child. We intuitively think otherwise only because most Indo-European langauges are similar to English in that they have a valid class "pronoun." We must get away from analyzing other

languages into classes on the basis of what they do or do not translate into Indo-European languages as.

To return now to the nature of languages, we have seen that there are two fundamentally different types of morphemes. So far as we know this is true for all languages, and these are major and minor morphemes. In highly inflective or agglutinative languages these are roughly stems and affixes, but this is only about 95% true even for these types of languages, and this dichotomy is not at all applicable to isolating languages, and probably not to synthetic languages, except to the extent that the synthetic or isolating language is also inflective or agglutinative. Since all, or nearly all, languages appear to be at least partially inflective, agglutinative, isolating and synthetic, minor morphemes will be affixes to the extent that a language is inflective or agglutinative, but they will be free forms to the extent that they are isolating. I am not sure just what the relationship with synthesis is at the moment. Now, the major morphemes appear to be fuzzy sets, just as we saw that vowels on the phonological level are fuzzy sets. The next question is, are minor morphemes possibly crisp sets which could be distinguished in terms of distinctive semantic features, i.e., semantic components?

In answer to the question posed at the end of the paragraph above, the answer has to be that at least some are. We are constantly describing such things as pronouns in terms of semantic distinctive features. For example, in English our pronouns have the following semantic distinctive features.

I	1st Person Singular	We	1st Per. Pl.
You	2nd Person		
He	3rd Person Singular, Masculine		
She	3rd Person Singular, Feminine		
It	3rd Person Singular, Neuter		
They	3rd Person Plural		

Note that there is only one "you" given above. This is extremely important, because setting up a you which is singular and a you which is plural is, again, analyzing English on the basis of what certain forms translate into in some other language. The fact about English, which is extremely important, is that as speakers of the language and as listeners who hear the language and try to interpret something that has been said to us, we cannot distinguish singular from plural in the second person unless there is some additional information in the stream of speech, the situation, or prior knowledge which tells us that you is representing a group of people. What you says to a listener is that the subject of that sentence is a group (possibly a group of one) and that that group includes the listener, but it may well include many other people as well.

Now, contrast the above set with those pronouns for Turkish.

Ben	1st Person Singular	Biz	Plural
Sen	2nd Person Informal	Siz	Formal
O	3rd Person Singular	Onlar	Plural

Note that here too there is no dichotomy between singular and plural on the second person level. The situation here has been described earlier, but just to refresh your memories, the dichotomy is between formal and informal, whereas English makes no distinction at all. One uses, in Turkish, the formal form to symbolize any group which might have someone who is the speaker's equal or superior in it, but he uses the informal, sen, for groups when it is clear that every one in the group is inferior to the speaker. Japanese, on the other hand, clearly distinguishes between singular and plural. The point here, again, is that when we talk about all languages in terms of singular and plural, 1st, 2nd and 3rd persons, we obscure some very very important differences between languages.

We stated earlier that when we make statements such as, all languages have pronouns, we also obscure some vital information about the differences between languages. Here we see the same type of problems.

I am not sure at this point that all minor morphemes are such crisp sets, but it is clear that at least some are, such as the examples given above, i.e., the pronouns in both English and Turkish. It will be noted by anyone who knows anything about Turkish that the forms I gave are, indeed, not pronouns at all, but are nouns, and in fact this was pointed out earlier in this book. However, those forms are easier to cite, and they correspond exactly, in terms of meaning, with the Turkish affixes which are true pronouns. The Turkish pronouns, as well as the English pronouns, are sets of forms which absolutely replace each other in all syntactic positions to express the semantic distinctive features listed above. These differ fundamentally from the classes of major morphemes, which generally, say 90% of the time replace each other in the same types of syntactic positions, and probably always could, but when two morphemes contradict each other in meaning, they are not likely to follow one another, though, Chomsky to the contrary not withstanding, when I was in the army and stood guard for 24 hours, after I was released, I "slept furiously." I think that any English speaker knows exactly what I mean, and that is a grammatical and meaningful utterance. By definition then, any grammar which excluded this from the possible utterances in a language would be a defective grammar, and we all know that if we have seriuosly thought about it.

It will be necessary to illustrate a number of paradigms of minor morphemes, because I know that speakers of English at least will misinterpret the class because of preconceived ideas. For example, it will come as a rude shock to many

that forms of the so-called irregular verb "to-be" are absolutely not verb forms. There is and never should have been any doubt of this. If for no other reason English teachers should have figured that there was "something rotten in Denmark," to quote a very famous English writer, from the kinds of errors that foreign students make when they are learning the language. For example, almost all foreign students are taught that verbs are predicates of sentences in such sentences as "I go." Then they are taught, "I am." Then when they are taught the question forms, they learn to say, "Am I." Then on their own initiative, they create "Go I?" If we stop for a moment and think about it, this is exactly what traditional grammar teaches them, and we have no one to blame but ourselves. If we kept forms of "to be" separate from verbs, as they should be, because they are in no way comparable with verbs, this would not happen. These are a small set of morphemes which replace each other under clearly describable conditions. Such sets should never, never be confused with classes of major morphemes, such as verbs or nouns, even though we know now that English doesn't have those two classes of major morphemes.

The forms of "to be" are listed below with their allomorphs and semantic components.

am replaced in alternation by is replaced in alternation with are
The above morpheme means non-past time.
was replaced in alternation by were
The above morpheme means past time.
been
The above morpheme is time free and is utilized when the morpheme is not functioning as the predicate, with forms of "to have."
be
The above morpheme could easily be called the infinitive form, as that is its most common use in English. It too is time free and functions not as the predicate.

It is extremely difficult to give a meaning to either this paradigm or to the elements within the paradigm. This is partially a result of the nature of minor morphemes. It is much easier to tell what they do than to tell what they mean, and this itself should tell us something about the nature of languages. The forms of be equate the subject of an utterance with the complement, if it is expressed, with the complement in a previously expressed sentence, or at times an understood complement. Note in the examples below, the only common element of meaning is that of equating the subject with the complement in some way.

I am tall.
I am Bill.
I am American
I am going.

I know that it will surprise many readers that I show going as the same type of element as tall, Bill and American, but this is the way one must view English, if he is ever to understand "How the Language Really Works." The point is that, tall is a modifier, and in the last sentence, going is a gerund, and one of the things that gerunds do is act like modifiers. Consider the man who was running fell and hurt himself and the running man fell and hurt himself. This brings us back to one of the fundamental points we have been emphasizing ever since we began, and that is that if there is a distinction between certain classes, there will be many places in the use of the language where that distinction will make it impossible for the reader to understand a sentence if he does not know whether a given word is a member of class A or B. The point above is that if one does not know whether the form is a modifier or a noun-verb, he cannot properly tell whether the person IS the thing or HAS the quality. On the other hand, a knowledge that table is a noun, is of absolutely no value to anyone who is trying to understand English, be-

cause we can tell from the sequence of morphemes in which it is found whether it is functioning as a predicate or something else, and that is all of the information that is relevant concerning table. We contend that this implies that the native speaker does not distinguish nouns from verbs, but must distinguish modifiers from nouns in his brain. We shall return later to the nature of gerunds, because this is really quite different from that pictured in most grammar books on English.

I need not go into it here, but if we consider the forms of "to do" and "to have," we find out that they too are minor morphemes and should in no way ever be associated in the minds of prospective speakers of the language with major morphemes such as irregular verbs. Perhaps you wonder how we associate irregular verbs with regular verbs. In order to put these in the same class, we need to find some frame into which the irregulars go in the same form as the regulars. This frame is:

He _____s (it). as in:
He eats it. He walks it. He types it. etc.

Note that in the third person everything that is called a verb takes the "s" suffix, but the forms of "have" and "be" do not. Forms of "do" also follow this distribution, but because one can say "Does he?" and not "Goes he?" we must conclude that this does not belong to the class verbs. This is not the only bit of evidence for such a conclusion, but it will suffice for the moment to illustrate the point.

It is necessary here to give a few more examples of minor morphemes, because these are difficult to handle, and one of the major problems in English grammar historically is that they have been grossly mishandled. Take the question-words as examples, these words have been handled in a variety of ways by both linguists and traditional grammarians, and all of these ways have been incorrect, unless I

have missed one somewhere, which is always possible, since there are literally hundreds of books that have been written about English grammar, almost all on an intuitive basis and all about equally useless to anyone who wants to use the language effectively.

The question words fall into two groups, each of which function quite differently in different situations in English. Of course, they have been lumped together by some scholars, because they do perform certain of the same functions, for example:

Who? When? Why? Where? Which? How?
What? Whom? Whose?

However, the differences become clear very quickly, when we put these into sentences.

Who is going?
What is going?
Which is going?
Whose is going? but

When is he going?
Where is he going?
Why is he going?
How is he going?

One can easily see that the second group cannot function as the subjects of sentences or clauses. They must have a subject in the sentence. Hence we see the form "he" in those sentences. Are these question words? Perhaps, but we note that the question is signalled to the listener in a sentence by the inversion of the form of "to be" and the subject. Note here that I did not say subject-verb inversion, because one thing that is absolutely impossible in English is subject-verb inversion. We get subject inversion only with forms of the three paradigms of minor morphemes, "to be," "to do," and "to have," but never with verbs. Another major error in traditional grammars of English.

What then do these morphemes mean? Below they are set into paradigms with a meaning for each given opposite it.

Where	at what place
When	at what time
Why	for what reason
How	in what manner

This is a paradigm of non-relative question words, because they cannot function in English sentences as the object of a preceding predicate and subject of a following predicate at the same time. They must always be followed in some manner by a subject for the following predicate. They can, however, function as the object or complement of a preceding predicate. These contrast with the following paradigm in that this next paradigm consists of relative question words, i.e., those question words which do, indeed, function both as the complement of a preceding predicate and as the subject of a following predicate.

Who	which person (whom is an allomorph)
What	which object
Whose	which possessed object
Which	which of a select group of objects

To give only a single example from each:

I don't know who went.
I don't know where he went.

Note the required subject following the member of the first paradigm above, i.e., where.

Since we are not concerned here with a complete analysis of English, but rather with the nature of languages generally, I shall refer the reader to my book "How English Really Works," for a further look at some of these problems and leave it at that. Basically we have spent some time on English because it is a good example of isolating languages and points up very well the problems encountered

while trying to deal with isolating languages. As stated earlier, this stems solely from the historical fact that our methods, theories and concepts of analysis have all derived from studies of highly inflective languages, and it is not surprising to find that the methods work much better for Turkish or Greek than for English and Chinese.

Briefly, we have seen that not only do languages not all have the same classes of major morphemes, but what is included in a given class, e.g., nouns, is so vastly different from language to language, that a statement that all languages have nouns is grossly misleading, when we are trying to find out how languages are the same and how they are different. For example, we saw that the so-called pronouns in Japanese are actually nouns. We saw in Turkish that they have two classes of pronoun-like morphemes, but that one paradigms is a set of minor morphemes but the others are simple nouns. English, on the other hand, has only minor morphemes to express the relationship of the speaker and listener to the predicate of a sentence or the nouns being possessed. All languages show possession in some way and all languages show relationships between the speaker, listener and others to the predicates of sentences. That would be a linguistic universal. However, the languages of the world do this in such different ways, that I am not sure we can make a great deal out of that bit of information without getting so far from reality that the description is useless, but we shall worry about that later.

All languages do have large classes of words which basically label classes of things that the members of the speech community want to talk about. We do not say "reality" here, because that gets us involved with whether ghosts and goblins are real or not. Whether real or not, these things need talking about, especially in some cultural groups. All languages also have paradigms of morphemes which have purely grammatical meaning, that is, they do not label things in the world, real or otherwise. The question words listed

above are an example of these paradigms. The classes of labels for significant classes of cultural items are fuzzy and they are major morphemes, and they appear to exist in all languages. The small classes are minor morphemes and also occur in all languages, and it appears that at least a large number of these are crisp. The examples that we have cited are the pronouns and the question words. These paradigms, and many others (though perhaps not all) are crisp sets each delineated from the other members of the paradigm (set) by some distinctive feature of meaning. The paradigm is delineated from other paradigms by the fact that it has a unique set of distributions. Here we see a striking parallel with the phonological level. We see a group of fuzzy sets (phonologically these were the vowels) and crisp sets (consonants). I believe that this dichotomy does, indeed, reflect something of the wiring of the human brain. The rigidity of the crisp sets is absolutely necessary for communication to take place, and without it, human communication would not be possible. On the other hand, the flexibility built into the fuzzy sets makes language the dynamic living vehicle for art and literature that it is. If it did work on the basis of crisp sets alone, it would become very stultifying and man could not have good poetry, or even prose. He would be just like a computer. Fortunately, on both levels looked at so far, we find that some of the sets appear to lend themselves to fuzzy analysis and some lend themselves to crisp analysis.

Not to labor the point too much, I will give a few types of morphemes which should be looked at as minor and not major morphemes. Some of the classes of morphemes that appear to be minor morphemes in most languages are: conjunctions, pronouns, question forming particles, all nominal inflections and all derivative morphemes, prepositions, and articles. Note that the distinctive characteristics of these morphemes are:

(1) small classes, usually less than a dozen, (2) they have a grammatical meaning, i.e., they do not label classes of objects in the cultural world of the speakers but rather show grammatical relationships, (3) many of them appear to be crisp sets. Major morphemes can be distinguished from minor morphemes because they: (1) occur in very large classes, usually hundreds of morphemes, (2) some, but not all of them are stems, (3) they label culturally significant classes, (4) these appear to be fuzzy sets. Examples of major morpheme classes would be the class noun-verb in English as well as the class modifier, and in Turkish the class noun and the class verb.

Now, we have set up a situation wherein morphemes in languages are set up into classes. For the moment, let us turn our attention to the major morphemes. As we all know, we often want to use a noun as a verb, or a verb as a noun, but if you speak Turkish, you simply cannot utilize a verb stem as a noun or a noun stem as a verb. This rule is certainly better than 90% accurate, and that is probably about as good a grammatical rule as one can come up with, because of the very nature of langauge, but that is considerable better than some of the present rules for English grammar which are accurate far less than half of the time. All languages have some method of creating members of one class from members of another class, and this process is called derivation.

Derivation is the process of creating "words" out of morphemes. One of the basic problems historically with understanding the nature of languages has been the failure to recognize that there is a fundamental differnce between morpheme classes and word classes. There are certainly in English, and I suspect in all languages, word-classes which are not parallel with any morpheme classes, i.e., there are words that fit into syntactic slots into which no single morpheme will go. We shall come back to this later, but for now, we must concern ourselves briefly with the process of creating words in one class from morphemes in another.

Of course, we all recognize that there is derivation without shift in class. For example:

Turkish	English Equivalent
göz	eye
gözlük	glasses
gözlükcü	optometrist.

In the examples above we see how Turkish has created new nouns from a single old one. This process can be extended considerably further than this in Turkish, but the above examples show the creation of a new word from an old one by the addition of a derivative suffix. So we see that we can alter meaning or shift classes, as shown below:

Turkish	English Equivalent	
boş	empty	(noun, to be empty)
boşal	divorce.	(verb, to divorce)

In the former case we have altered the meaning but not the class. In the second we have altered both the class and the meaning. However, so far as derivation is concerned as a general process in all languages, it is always necessary to alter one of the other, but not necessarily both. However, it quite often is a fact that both are altered, as in the second case above.

To show how this process operates in English, we must look at two morphemes which have been traditionally labelled as inflective, but which are clearly not inflection. These are:

-ing
-ed

We are told in grammar classes that the former is a progressive inflection, if we are told anything at all, and that the second is a past tense inflection. However, when in Singapore working at the Southeast Asian Ministers' of Education Regional English Language Center, I did a careful count of uses of this form, and I found that noun-verb stems with -ed affixed, symbolized past time less than 40% of the time. The other occurrences were time free.

It is absolutely clear that the function of these two morphemes is something very different in English from what we have been taught. What exactly do these two morphemes do? Clearly, they are derivative morphemes and not inflective, because once one of these has been affixed to a stem the syntactic functions of the word so affixed are very different from the meaning or use of the stem before such affixation. For example,

Running to school he fell;
Burned badly, he died in the hospital.

In the two examples above the forms running and burned fill a slot which cannot be filled by any single morpheme. Hence we have indication that we have created a "word-class" which is different from any "morpheme-class" that we have in English. Further.

The rolling stone hit the car;
The burned building collapsed.

In the examples above both rolling and burned are performing a modifier function, not one which could normally be performed by either a noun or a verb. Further,

His comings and goings were watched carefully.

In this example we see that the -ing forms can both be possessed as well as pluralized, and this is the realm of the noun, so clearly we have done something to the noun-verb roll which makes it appear to be something like a pure-noun at this point. Can it function as a predicate for a sentence any longer? Absolutely not!!! The -ing forms cannot be the predicates of sentences, and as a result, in sentences such as I am going, we have to have a form of to be with them to act as the predicate.

What then is the -ing form? We might just as well call it a gerund as anything else, because

that is what it has always been called. Now all we have to do is define a gerund in some sort of manner which is comprehensible to learners or speakers of English. This is remarkably easy. A GERUND IS A VERB-NOUN STEM WITH -ING AFFIXED TO IT. The forms are clearly identifiable. Any person can be taught to recognize one in a matter of minutes. Further, I believe that since the form is what the listener hears and the speaker produces, this parallels exactly what goes on inside his head as he hears or speaks the language. What do gerunds do in English? This too is somewhat different from what we have always been taught. GERUNDS PERFORM THE NORMAL FUNCTIONS OF MODIFIERS, ACT AS VERBAL COMPLEMENTS, FUNCTION AS PURE-NOUNS AND FINALLY FILL A SYNTACTIC SLOT WHICH CAN ONLY BE FILLED BY GERUNDS OR PARTICIPLES. We will leave the definition of a participle for a little later, but it too will be clearly and easily identifiable and have clearly statable syntactic functions. Which is very different from the grammar that is taught today for English.

It is necessary here to illustrate some basic fallacies in present English grammar to show the importance of these new distinctions. In sentences like the following one, we find that traditional grammar teaches us that the -ing form is the predicate, but this is impossible. Note below.

The man was running and fell down.
The running man fell down.

Obviously the word running is performing the same type of function in both sentences, that is, it modifies or describes the subject, man. Hence, it could not be the predicate. Further, the tense is carried by the predicate. What carries the tense when a form of to be is present? The forms of to be. Note the following examples:

Was the man running?
Yes he was, and fell down.

There is tons of additional evidence on this score, but since we are not interested in analyzing English here, but rather showing the nature of natural languages, we will leave this. Suffice it to say, that it was necessary to point up some of the short-comings of English grammar to show the relationship of traditional grammatical methods to the errors in description which blind us to much of the natural character of linguistic systems.

Now, for a brief look at participles. What is a participle? A PARTICIPLE IS A NOUN-VERB STEM WITH THE SUFFIX -ED AFFIXED TO IT, OR IT IS A SPECIAL, IRREGULAR, FORM WHICH PERFORMS EXACTLY THE SAME FUNCTIONS AS THE SO-CALLED PAST TENSE FORMS OF REGULAR VERBS PERFORM. Here again we see that now we can teach anyone who wants to learn English exactly what a participle is, and we are not constantly running into the problem with students saying, is that a participle or a gerund. These two elements in the grammatical structure of English are very different, grounded directly in the forms which the listener hears or the speaker produces, not some amorphous meaning which no two teachers of English grammar can agree on. We are now ready to give the syntactic functions performed by participles, at least tentatively, though this has not been carefully and fully analyzed as yet by any means. PARTICIPLES PERFORM THE PREDICATE FUNCTION AND WHEN SO DOING MEAN PAST TIME. THEY ALSO PERFORM MODIFIER FUNCTIONS AND WHEN SO DOING ARE TIME FREE, AS WELL AS THE SINGLE FUNCTION DESCRIBED FOR GERUNDS WHICH CAN BE PERFORMED ONLY BY PARTICIPLES OR GERUNDS.

What have we been doing above? We are demonstrating the fact that English has at least two classes which are derived with the forms -ed and -ing which do not correspond to any morpheme

classes in the language.

Not to belabor the problems of English grammar but, rather, to illustrate the problems in dealing with isolating langauges, let us look briefly at the difference between the morpheme classes and word classes.

Morpheme Classes	Word Classes
Noun-verb	Noun-verb
Modifier (includes adverbs)	Modifier
Particles	Particles
(all minor morpheme paradigms)	Gerunds
	Participles
	Pure Nouns
	Pure Verbs
	(some, but not all classes of minor morphemes)

Note that all of the morpheme classes (except a few minor morpheme classes) are also word classes. This will be true to the extent that a language is isolating, because all of the stem morphemes, such as, noun-verb, function as words when they have no affixes attached to them. However, note that there are a number of word classes which do not really exist on the morpheme level. Examples of each of the classes are: run (noun-verb), but running (gerund), ran (participle, irregular form), runner (pure noun). Here we have to shift to another stem, because run will not form a pure verb on the word level, but sharp (modifier) will, in the form, sharpen (pure verb). This does not begin to show the intricacies of English grammar, but it should demonstrate the point that word classes and morpheme classes are not the same thing, and that when studying a language one must know the difference and describe both if we ever hope to understand how languages work. Certainly no definitive analysis of English has been made, either by myself or another scholar or combination of scholars. However, what we see above is the way one must approach a reanaly-

sis in order to see the true nature of the structure of this language.

What we see in the traditional descriptions of English structure, both from the linguists who have worked on it and those Latin grammarians, because in both cases we are dealing with scholars who have been so brain-washed through years of the educational process that they could not see the forest for the trees, is a set of erroneous classes, based on how things translate into other languages, and sets of rules which supposedly tell us how to use the morphemes in each class. However, since the classes of morphemes are different from the classes of words, and only one set of rules are given, the rules cannot be true more than a small percentage of the time. Further, since the original classification is not correct, rules do not work to help us make correct and appropriate sentences. What we see through the reanalysis of morphemes into the classes: noun-verb, modifier and particle, is that we now have a proper set of classes which reflect what the speaker utilizes when he speaks and what the listener has to analyze when he hears. With morphemes in their proper classes, we are in a position to create (or discover) rules which will enable us to help people who want to know how English works. Further, we now see that on the word level, we are dealing with a different set of classes, and logically it follows that the same set of rules will not work for words. We also see how much that has been traditionally called inflection is, indeed, derivation, and we see how simple the classification system really is and how one shifts morphemes from one word class into another when he needs to do so. Over the years what needs to be done is to continue reshuffling the morphemes until all of the morpheme classes are discovered. The rules need to be restated until they work about 95% of the time, and we will then have a really accurate and useful grammar of English

Now, what does all of this tell us about the

nature of natural languages? It tells us that a large number of language analyses do not reflect the true nature of the languages, because no one has bothered to seperate out the word classes from the morpheme classes. It tells us that isolating languages, and quite probably all languages, have two distinct systems, one which covers the way morphemes work in the language and one which covers the way words work in the language. It should surprise no one that we come to this solution, because if we look back through history, we see that the Greek and Latin grammarians saw all of language as a single system. As more was learned about the nature of more languages, we began to discuss "phonology" and "grammar," that is two levels of analysis. As time passed and more information was garnered on more languages, linguists recognized three levels, i.e., "phonology," "grammar" and "syntax." We then saw the introduction of "distinctive feature analysis" and the dropping of "phonetic analysis" from the field. Now, what we see is a combining of all phonology into one level, but an accurate description of that level of analysis involves both distinctive features and fuzzy sets. Then above that we have a system which consists of classes of morphemes and a set of rules for combining these into words. Then we have a system of words which consists of a set of word classes and a set of rules which tells us how to combine those classes. In the next chapter we shall take a brief look at how languages are learned, because much has been made in some circles about the importance of language acquisition as a determinfactor in creating our picture of the basic wiring within the human brain which makes human speech possible. Then, beyond that we will take up those levels above that of the word, i.e., what has traditionally been called syntax.

CHAPTER IV

In order to discuss language learning, we need to consider some body of data on children learning languages (cf. *The Development of a Linguistic System in English Speaking American Children* and its companion volume *The Development of a Phonological system in English Speaking American Children* by Joe E. Pierce and Inga VanWaardenburg Hanna, The HaPi Press, 1978 and 1974 respectively for a full report on these samples). The sample under discussion in the above mentioned sources is a collection of children 50 at each of a number of age-levels, 25 boys and 25 girls, ranging from 3 months of age through 18 months of age. While this does not cover the entire range of language acquisition, it does cover much of the formative period. However, even from this short run, covering the speech habits of 250 children, we can make some fairly good descriptions of what probably goes on in the mind of the child as he acquires his new language.

The fact that the children in the above mentioned samples perform at a level which is far below that reported by most published studies on eighteen-month-olds cannot escape anyone who has scoured the literature in search of norms for any age group. There are a number of possible explanations for this, but we believe that the real explanation is a very simple one, that of sample selection. Our sample, for the first time in history, is a thoroughly representative sample of normal children in the Portland, Oregon population. The sample was matched to the general population on the basis of a number of variables which are usually associated with levels of verbal performance, such as, economic status of the parents, educational level of parents, and so on. This has never been done

before. In fact, the studies which have been re-reported extensively in the literature all have been made on extremely small samples, which are clearly not average or representative of any population as a whole. For example, Jakobson's report on the Royal family hardly represents anything, except a single family, and certainly an unusual one. In Besançon, France, in 1974, a conference was held on *Prelanguage*. At this conference, hours were consumed by speech therapists giving reports on the development of their own children. Very few of the samples published actually look at more than three children and those that do rarely took a serious look at more than seven or eight. These could hardly be expected to be representative of the population generally, especially when virtually all of the children were either the off-spring of linguists, langauge teachers or speech therapists. These are highly verbal families, and their children would be expected to perform well above the norms for the population as a whole. We cannot extrapolate too much and say that the sample described above represents all English speaking children, or even all English speaking American children, but it certainly is a better sample than any studied heretofore. We would say that we believe that it fairly represents the general run of the Portland population and leave it at that. Those who wish to can draw what inferences they like for the population as a whole, at least until another significant sample can be drawn and studied to be compared with this one.

At this point we would like to ask, what conclusions can be drawn from the data published in the above mentioned study? In order to answer this, we need to go back to the beginning, i.e., the three month sample. At three months the average child produced a relatively small number of sounds or perhaps more accurately, a relatively small range of sound, thought one might say that there was an infinite number of sounds within that range. All of the vowel sounds found for the average child were

in a small oval extending from about the middle of the low-front region of the vowel chart up into the mid-central area and from about the middle of the low-central up into the mid-front region. A few of the three-month-old children produced a very low high-front vowel, no low back vowels at all, a few high-central vowels and a very very few very front and very low high-back vowels. The average child also produced three sounds that could be called consonant sounds. This is a far cry from the "infinite" number of sounds reported to be produced by small children in much of the linguistic literature. As we moved up to the six-month-old sample, we saw the range of vowel qualities double, and the number of consonants increase from three to seven, here too almost doubling. By nine months these figures had just about doubled again, and by the end of the first year, i.e., at twelve months, the children were porducing, on the average, all of the vowel qualities represented on the usual vowel chart, but with no distinction at all between rounded and unrounded vowel qualities. The development in consontant inventory was highly individualistic, in that each child learned a different set of consonants, but the total number appeared to increase at a regular rate. Less than forty percent (40%) of the twelve month old sample knew even a single word, so the typical twelve-month-old child in our sample had not yet crossed the threshold from making playful noises to connecting these noises with something in the real world.

Is what we see above representative of a learning process or of a developmental process? We feel that it must be viewed as both. For example, it seems very clear that the increases in number of phonological units controlled by each child must represent some sort of maturational process. However, the fact that each child produced different sets of sounds at each age level would seem to indicate fairly clearly that these are learned, either by accident or by an attempt

to produce what those around them produce. Virtually all of the sounds produced by these children fall well within the range of sounds produced by adult English speakers either as allophones of phonemes in English or as the basic phonemes themselves.

The first criticism that one hears when he begins to speak about child language before a group of linguists is that you are associating the sounds of the children with the sounds of adult English speakers. We made a very strong conscious effort not to do this. The purpose of charting areas of vowel quality on charts was to get away from English phonemes. We plotted the approximate tongue position for each sound on a chart and then drew a circle, oval, or whatever around the area covered by the dots and lines which represented the sounds for a single child. Then we combined all of these onto a single chart, and this gave us ranges of vocalic quality, not represenatives of adult English phonemes. In terms of consonants, we attempted to use a symbolization system which represented the articulation of the child so as to keep this separate from any connection with adult English consonants.

A combination of the way the data was collected and the way the results came out lead us to believe that the child learns the overwhelming number of sounds produced from those around him or her. If this were not the case, there would be a lot more non-English vowel sounds produced by the children in the sample. For example, there were no strongly rounded front vowels produced by any of the children, even at twelve months of age. In fact, the only non-English sound that occurred more than once was the bilabial fricative, "ɓ." From watching the mothers with the children, it seems clear that every time the child made a sound which by the wildest stretch of the imagination could be interpreted as a word, the parents reacted with very strong positive reinforcement. When a testor or parent asked the child to make any kind of noise, either a word

or repetition of a nonsense syllable, and the child produced a noise that even remotely resembled what was asked, the parents would light up as if they were electric light bulbs, often pick the child up if this were possible and in general heap positive approval upon him. When he failed, the least in the line of negative reinforcement that occurred was a slight frown, and some parents actually scolded the children for not performing properly. The only part of language that appears to be controlled by maturational processes from these data are the rough numbers of units and so far as articulation is concerned, the general area in which the articulation took place and the types of articulation.

By eighteen months it appears that all children can distinguish between consonants and vowels and generally produce the basic vowel qualities of English. The children could distinguish stop consonants from non-stop consonants on the average, but little more than that so far as consonants are concerned. Some could distinguish three types of stops, and at least one could distinguish voiced from voiceless stops, but most of the children could not. The average child could distinguish stops from non-stops and make at least one distinction between either bilabial and dental stops or between stops and nasals as well as semivowels or fricatives, but no more. The average eighteen-month-old child still has a very long way to go before he has even a simple grasp of the consonants used by English speaking adults. With an average of only three or four consonants, it is difficult to assume that the child has more than the remotest grasp of what consonants are and the basic structure into which he will eventually fit all such units. Even the two or three most complete systems controlled less than half of the English consonants and produced very little in the line of sounds that were not normal allophones for adult English consonants.

Now that we are at the eighteen month level, we can see the general outline of the way language structures are built up in the mind of a child. In the first year there is the maturational process which makes possible the production of a fairly wide range of noises. At twelve months, however, the child has no idea of a structure, since the average child has not as yet learned his first word. By eighteen months the average child knows a few words but as a rule uses all of them as single-word, whole-utterances. Hence the only structure that the child appears to be able to handle is phonological at eighteen months.

By looking at the eighteen month sample cross-sectionally, however, we can see that the children with the least linguistic ability have only a single word or two. Those with the greatest range of linguistic ability have several classes of words which they put together following clear-cut rules. This shows that they have acquired or learned a structure in some way. By looking at the nature of these simple structures, we believe that we can see the way linguistic structures develop throughout a person's life, only in microcosm.

Most of the children have only a single class of words all of which are simple labels for something in their environment that they have reason to designate, such as, daddy. We use this example because it was the one word that all of our children who used at least a single word, used. This word for some children labelled their biological parent only, for some all adult males and for others all people. There was no consistency at all in terms of the words learned by the children. This indicates clearly that the specific words acquired were learned from those around them.

Some interesting examples of the learning process should be examined here. For example, the children appear to learn a word like daddy associated with their father. This is the most common first word for the children in our sample, and

this is probably so because the mother is taking care of the baby and father comes in and out of the room often, and as a result the mother keeps saying something like "here's daddy," or "here comes daddy." The constant repetition of the sequence daddy even though it is embedded in a longer utterance is learned. For awhile, the child utilizes this label to symbolize only the thing that the stream of noise was first associated with. Then he extrapolates from this meaning, as in the case of the children who utilized the word daddy for all adult males or at least some collection of people including the father. Other examples of this phenomenon are children who used the word cake for all solid food and coke for all liquids. A child used milk for all foods both liquid and solid. There are dozens of examples, but they all add up to a child learning a word associated with a specific item or act and later extrapolating to other items or acts on the basis of something which we cannot always grasp. This process of extrapolation appears to be highly individualistic and not a part of the maturational or logical process of language learning.

When I (Pierce) was much younger than I am now, my own child knew only a single word at eighteen months. The word was "whazzat," which she used for "what is that?" I thought that she was feeble minded because I had expected her to be able to say a lot more than that by that age. On the basis of this sample, she was about average. She could not be below average in intelligence, because her college record was outstanding and she now has a PhD, so her apparently slow start did not effect her later linguistic development. This process of learning and extrapolation is illustrated clearly in an incident with her when she was less than two years old. She and her mother went out the front door of our house-trailer, and she saw a group of dogs. Her mother said "doggie." She repeated, "doggie" much to our delight, because she had never said that word

before. Then we went for a ride in the country, and down the road she pointed to a cow and said, "doggie." Her mother corrected her. She frowned but apparently made a mental note of the affair, because a little farther down the road she saw a horse and said, "cow." This appears to be the process of generalization, and here because we have more particulars, we can at least make some educated guesses about what went on in the mind of the child. First, she saw a strange animal which her parents labelled a doggie. She saw the general shape and formulated a hypothesis, all animals shaped like that are doggies, not in words, of course, since she had very few of those at her disposal. Then, the first time she saw another such animal, she tested this hypothesis by calling the cow a doggie. When this did not get the desired results from her parents, she concluded that small ones were dogs and large ones were cows. She tested this hypothesis, just as she had the first one, as soon as she saw another large one, and it too turned out to be wrong. The animal was a horse. It would seem that this is typical of the way children learn new vocabulary items, associate them with some collection of items in the real world, and finally refine them until their classification approximates the range of meaning utilized by the adults around them. The wide ranges of meaning used by different children for each word indicate that the children learned the word, generalized on the basis of his or her personal perceptions of whatever he or she saw and finally by trial and error associated with positive and negative conditioning reduced the overgeneralization in such a way that their categorization approximated that of the adults. Later in life, this process is dominated by their agemates, but at eighteen months, it is still controlled by the reactions of their parents, i.e., when they use the word correctly they get both what they want and praise. When they use it incorrectly, they often do not get what they want and sometimes are punished.

What do we need to assume in order to explain the observable data in terms of "deep structure" or the biological wiring in the human brain? All that is

necessary as an explanation of what is seen in this data is the ability of a child to associate a stream of noise with something in his personal perceptual field, i.e., something that he can see, hear or sense in some other way. Second, we need assume the existence of a storage device which will hold the electrochemical output of the child's sensors, i.e., the visual images probably represent a constellation of electrochemical outputs, probably in the form of a constellation of positive and negative charges. In this storage section of the brain there must be the capacity to store one constellation of charges which represent the aural output from the hearing device in one spot with the output of parallel "other" sensors opposite it so to speak. This device must be such that a very large number of physical, non-aural outputs can be stored opposite a single sequence of sound outputs. All this means in simple English is that the child must be able to retain the connection between the stream of noise coming in through his ears and the "other" sensors, such as the eyes, in a manner such that these will always be associated in his mind. This explains the facts that the child hears a stream of noise, say "bo" associated with his favorite toy. Then for a few days he uses the verbal symbol only for that toy. Then he hears someone use this symbol to refer to some other toy, perhaps a ball in his room that is a different size or color. On the basis of a highly individualistic perception of the situation, the child then begins to store all "similar," that is, what he considers to be similar, items on the basis of the criteria that he has set up as distinctive features for that class. This same procedure appears to be followed with each item learned. Then, he hears someone call a <u>bowl</u> something that sounds similar to him, and so he stores the mental picture of a <u>bowl</u> opposite the same symbol as he has stored <u>ball</u>. For a time he gets along using the one symbol for both items, but soon there comes a

time when he wants a bowl, asks for it and someone gives him his ball. He is hungry and does not want the ball. This is strong negative conditioning, and he realizes that there must be some difference between bowls and balls. At that point he listens carefully to hear some difference that he can use to distinguish his one want from the other one. It is this process which makes the child identify the phonemes. For some sounds, particularly stop and fricative consonants, this is probably a distinctive feature. For vowel sounds and the /r/ and /l/ like sounds, it is probably not so. These latter phonemes are probably fuzzy sets, as discussed in the chapter on phonology, and these are identified as described there.

It seems to me that only within this type of theoretical framework can one explain facts such as the variation between /ɨ/ and /e/ before /n/ in my dialect, as discussed earlier. The fact that one child appears to distinguish between the words ball and cat on the basis of the vowel and another one distinguishes the two words on the basis of the initial consonant is certainly not explainable in any other terms that I can imagine. What I mean is that one child will have a word that sounds something like /gæ/ for cat and a word such as /bæ/ for ball. However, another child will have a word that sounds like /bæ/ for cat and /bo/ for ball. By no stretch of the imagination can distinctive feature theory handle such wide-ranging and idiosyncratic differences among children. For the same reasons one must give up any ideas that this is a developmental or maturational process. If it were maturational or developmental, there would be a great deal more consistency between children than we observed in our sample.

For the observable data, we need assume no more than we have given above. There is no need to think that the child thinks out a sentence in some sort of deep structure and then converts it into surface structure. The child merely sees, hears and feels the world about him. These pictorial images can be

stored in a spot in his brain. Opposite these images or sensations, he is capable of storing a stream of noises. As his linguistic system increases in number, i.e., as he learns more symbols (words or morphemes) he simply lines them up at first. Most children accumulate a store of about ten to twelve of these meaningful units before they note that some of these units can be used differently from others. This too needs no assumptions about thinking per se, only the ability to hear that one of the words he has been using as a whole utterance occurs in conjunction with another word he has been using as a whole utterance, and that when these occur together one precedes the other. When he begins to use these sequences, they appear to be mere repetitions of what he has heard and he appears to treat them as if they too were simply new words. Then after hearing a few of these sentences containing the same word in initial position followed by other words that he knows in second position, he isolates each of the elements. When he has done this, he has learned the basic fact about a language system. From here on in, language learning consists merely of classifying the words and learning how he can and cannot put these classes together to form sentences.

Note the description of the speech of the eighteen-month-olds shows no consistency at all in the words learned, the way these are classified nor the rules for putting the classes together. This indicates quite clearly that there is no underlying "logic" in the system. The child hears utterances, at first he recognizes only single words. Then he notes that one of his words occurs in what he considers to be other words. Then he analyzes these complex words into sequences of two words. Then he at first puts all words that he can use in initial position into one class and those that he cannot in another class. Then he "creates" new sentences following the rule, members of class A always occur in initial position in sequences. Then he learns, by

hearing them in context, that some words can be found in both initial and final position. Here he has established a third class. Somewhere along here he has heard a few utterances in which one of his two-word utterances is connected to one of his one-word utterances to form a three-word utterance. In the eighteen-month-old sample this is as far as the process goes. However, recent research by other scholars has demonstrated four and five word utterances, but they report that there is no three-word stage. However, we have to insist that some of our sample clearly had a three-word stage in their acquisition process. Utilizing the types of utterances described above, the child tears them apart and rebuilds them to form longer and longer sentences. As he grows he continues to refine his classification, again on a completely individualistic basis, based on the particular sentences he hears the most from his associates, not on any logical or relational basis. In the data, not only do the number and types of classes found in the linguistic structures of the different children differ, but the elements in a particular class differ, e.g., two children have a class which we have labelled as "denotative," but in one we have where's and that's, whereas for another child a similar class includes the "words," here and that. The only linguistic features which could be developmental would be those that were shared by all of the children. Hence we do see at least one item which might be developmental, and that is the fact that the children all do appear to learn to utilize single word utterances first. Some of these could be analyzed by a linguist, but there is no indication that the child breaks them into units. First, these all belong to a single class which is utilized to symbolize a class of items in the real world around the child, but as time goes on, he learns to combine two of these into real two-word sentences. Then he learns that these words that he has been learning fit into classes. The specific classes are individualistic, hence not developmental, and the words in

the classes are individualistic, and hence also not developmental. The child learns at this point, or perhaps we should say realizes, that the entire basis of language is that there are classes of words and rules for sequencing these classes, NOT individual words, but classes to form sentences. Then he learns to add a third word, by combining some of his two-word sentences with some of his single-word sentences. Once this is done he recognizes also that there are more classes than he has used to that point. None of this is conscious thinking, incidentally, it is merely the recognition of words he already knows in sequence with other words he knows. He sees on the basis of the use of his parents that certain words, such as ball, cake, etc. seem to go together, so he stores them in the same place in his brain. There is probably nothing more complex about the concept of classification than that. The brain is so constructed that the child automatically stores those items that seem to appear in the same slot in a sentence in the same part of the brain. Once he has stored these in the same place, his brain will go to that place to select a word to go in initial position to form a new sentence, to the other place to find an item that goes into second place to find an item that goes into second position, and so on down the sentence for any length sentence he wants to put together. It is the programming of the search pattern in the computer-like brain which makes language possible, nothing more and nothing less. Hence language structure consists, on all levels of discourse, of classes of units, i.e., words, morphemes, phrases, etc., and rules for combining these elements into sequences to form proper utterances. The child gradually learns to extend the words from two, to three, to four, until he gets to the point that he sees sequences of two or three word utterances put together. Hence it is this differential learning process which accounts for the fact that different people consider different

things to be grammatical, not their poor performance as measured against an abstract system.

We have carried the concept of linguistic learning somewhat beyond what we have observed in the data, but it appears that this is merely the logical extension of what has been observed, and some researches performed by others have verified some of our ideas even though our sample would not. As to the vaunted "creative" abilities of man, one sees, if he takes the trouble to examine carefully the process in detail, that it consists merely of putting an element from a class common to a particular position in that position when no one has done so before, or occasionally breaking the normal rules for effect. Take a simple linguistic process such as that of creating nonsense words. If you ask native speakers of English to create such nonsense words, they will merely put together sequences of English phonemes in normal order, but ones that do not occur as regular words. One does not get sequences, such as, /pprss/, because this is, indeed, an impossible sequence for an English speaker. The initial sequence of two released stops is impossible, but not in other languages. If we were very creative we could get unique and different sequences of phonemes in such nonsense syllables or words, but we do not. The sequencing of phonemes follows the normal rules of English. The only exceptions that I have ever found in my linguistic classes have been when one of my students spoke some language other than English and he then put together words following the sequencial possibilities for that languages instead of English. What often passes for "creation" stems from the fact that a particular individual, because he has not been led down the garden path by traditional grammarians, realizes that a certain word, in his linguistic structure or dialect, is a member of a class when everyone else thinks it belongs to a different class. When he writes or speaks, he uses it as an element of the class to which it belongs in his speech. This does not constitute either poor performance in the language nor any

great creative ability. Winston Churchill was a good example of a man whose speech differed from that of the entrenched establishment in English, but he was recognized as a brilliant speaker, mostly because he used his language the way he learned it and did not allow the establishment to frighten him into dropping his own speech for an absurd norm which is logically derived from Latin but which has little or nothing to do with the way English should be used. He was able to get away with this because he was a powerful political leader. An ordinary mortal, speaking the same way, would have been graded by many as "incompetent."

To recapitulate briefly, we see from the data presented in this chapter, in the mind of a child at birth an inate ability to associate sound symbols with sensory inputs from his environment. Further, he associates a number of inputs with the same symbol and stores them together. Then he learns that only some of these really belong to that class, so he pulls out those that do not belong, as experienced in his daily life, and puts them into a different class, i.e., opposite a different symbol in a different place in the brain. Then he learns that these symbols form different classes, and that the classes can be combined in certain sequences and not in others. He learns to make one-word sentences, then two-word sentences, and so on until he can make extremely long drawn-out sentences. The learning of all of the words and their meaning is through a process of positive and negative conditioning, nothing more, as the person gets what he wants when he uses a word or sequence in a manner that transmits his wishes to someone else and does not get what he wants when he does not transmit a clear message.

What has been said above has certain logical implications for adult English. So far as we could observe, the child thought only in the words that he knew so far as language is concerned. He thinks in terms of the classes that his language provides.

It is this fact that led Sapir and Whorf to lean so heavily on the idea that language actually controls our daily behavior (cf. Benjamin Lee Whorf, *Language, Thought and Reality. Selected Writings of Benjamin Lee Whorf*, ed. John B. Carroll, Cambridge: Technology Press, 1956 and Edward Sapir, *Selected Writings of Edward Sapir in Language, Culture and Personality*, ed. David G. Mandelbaum, University of California Press, 1949). If your language does not have a neat little paradigm of question words, i.e., who, which, what, etc., you are not prone to ask those questions. This does not mean that the language is incapable of asking such questions, it means that such questions are not forced by the structure of the language a person uses because these words come first in sentences, and the oppositions are so clear that an English speaker can hardly keep from asking the questions even if he doesn't want to know the answer. A Turk does not ask them at all in many instances and Americans think that Turks are insensitive and so on. The fact is that their language structure does not condition the Turk to automatically ask such questions, and they will only do so if they have a vested interest in the person or thing that is being discussed. It would appear on the basis of this study that it is clear that much of our thinking is directly controlled by the words and structures that we have available to us with which to express our ideas. For example, English speakers rarely feel compelled to state the degree to which they know that an event happened, but Turks do it almost constantly, because they have a simple form which suffixed to a verb tells the listener that they are reporting information, i.e., the -mış verbal suffix. Is this not a control over behavior of sorts?

It is often said that "everyone wants to say the same things." This shows only the ignorance of the person making such a statement, because when you learn Japanese and later go to Japan, the first thing you note is that the grammatical structures you have learned are those that express the things

that English speakers want to say to each other, and that the Japanese seldom want to say any of those things. You also find that you do not know how to say the things that the Japanese want to say all the time. Cultures determine what one wants to say and the structure of his language gives a person the capacity to say these things. If there are any universals in language at all, other than the fact that they all structure in a certain way, it is extremely difficult to see them. To say that all languages have nouns or noun phrases is to make a meaningless utterance, because the class noun in English is certainly not comparable with the class noun in Turkish or Japanese, if, indeed, one can make a case for the existence of such a class in English at all. The class noun in Turkish contains quite different things, has different syntactic functions in the language and in many ways is so different from nouns in other languages mentioned above that it is a little like trying to compare cows with germs because they both have life. Further, if one does not have a class "noun," then it is difficult to have a noun-phrase.

Languages appear to be learned systems consisting of verbal symbols to which ranges of meaning in reality are attached. These symbols form classes and each language has rules for combining the classes that are a part of that structure into meaningful utterances. The only natural ability that a human being seems to have is that of classifying and generalizing, an ability that is utilized to its fullest at eighteen months of age. That is, this does not have to be learned in school. In fact education curtails this ability. It does not enhance it.

In this chapter we have attempted to give a justification for the study mentioned and drawn certain conclusions which appear to us to be reasonable on the basis of our observations of the 250 children studied. We have not attempted to prove or disprove any particular thoery, but rather

have "done our damndest to find the truth" as it flows from the study, and according to Bertrand Russell, that is the essence of science.

For more information on the relation between language and other forms of human behavior, cf. Joe E. Pierce, A *Theory of Language, Culture and Human Behavior*, The HaPi Press, 1972.

CHAPTER V

The last chapter was a brief digression, but it was felt to be a necessary digression, because of some of the conflicts that rage in linguistic circles these days. If we have a theory about the nature of deep structure, and if much of that theory is based on beliefs about how languages are learned, then we need to verify or reject certain of the ideas before we can move on to clarify our picture of the nature of language and the relationship between language and deep structure. Now it is time to move to those levels above that of the morpheme and the word.

What has usually been called syntax cannot be understood in terms of any of the existing theories because they attempt to study it as if it were a single ball of wax, so to speak, and it clearly is not. As we saw in our discussion of the morpheme and word "levels of analysis" or systems, since each constitutes a complete system, we only confuse the issue and obscure many important points about the nature of language, when we attempt to talk about them as if they were a single system. The same is true on the syntactic level. That this label covers a multitude of sins is clearly evidenced by our use of such labels as "phrase, clause, sentence," and so on. The picture of syntax is so unclear that linguists can argue about it until hell freezes over, if they wish to do so, because the picture will never clear up one iota until we realize that we are dealing with a complex interrelationship of a number of systems, not a single system, and that these systems work independently within the mind of the speaker as he is speaking or listening to the language in a conversation.

Let us now attempt to understand, at least a

little better, what kinds of levels there are in what we have been calling syntax and how these systems or levels work and how they interrelate to make human communication possible. First, from a variety of studies, the basic, or kernel sentence in English, and probably in all languages, is the single-word utterance. We see here another case wherein isolating languages are vastly different from highly inflective or agglutinative languages. They are different in that many things which are affixes, and by difinition cannot occur as whole utterances, are free forms in isolating languages. For example, in English one can say,

Did you say in or out?

and the answer is,

Out!

In Turkish you cannot do that, because Turkish is highly agglutinative, and the forms for in and out simply cannot be said except as a suffix on a stem. For example in Turkish one must say,

İçeride mi aşağıda mı?

and the answer is

Aşağıda. Literally in the outside.

The form aşaği means outside, and the da is an affix which means in that location.

In English, then, all morphemes, major and minor, appear to be utilized in normal speech by native speakers of the language as single-word whole-utterances. This is the basic or kernel sentence, because it is the simplest form of syntactic construction and also because that is the form of sentence first learned and utilized by children, so it makes a very logical starting point for our consideration of the nature of lan-

guages. Actually, this has been one of the most significant points in the descriptions of languages in the past. The very fact that words or morphemes were classified in terms of affixation means that the morphemes which were not words were affixes, and by definition could not be used as whole utterances. In English there is such a small list, that probably the best thing to say about utterances of this type is that all morphemes can occur as single-word whole-utterances, except, and list the handful of affixes in the language. Of course, by definition, all derived words, that is all classes such as gerunds and participles which are, indeed, sequences of morphemes, can also create this type of utterance. Does this happen in all languages? I don't know, but I suspect that it does and that this part of the structure has been obscured by traditional description.

Since there is not much to be said about the structure of single word utterances, let us move quickly on to the phrase structure. A lot has been said in the literature about the efficacy or lack thereof of phrase structure analysis. Much of the criticism is valid, in that once one has made a phrase-level analysis of a language, he has only described a small portion of the structure, and one needs to go on and finish the job. However, it is equally clear that one gains very little if he simply goes up to the sentence level and describes that system, because we are dealing with a hierarchy of systems each of which is a basic cog in the next higher level. One of the main problems is that we do not actually know how many levels exist in any given language and whether or not any given level operates in all languages, or if each language has a unique number of levels. These are some of the things that need very badly to be explored seriously by linguists.

What then is a phrase, since that is clearly the level immediately above that of the word? The phrase is a unit higher than that of the

word, i.e., some sort of sequences of words. What we need, then, is some way to demonstrate that certain types of word sequences operate as units on a higher level, and this is not too difficult to do. For example, if we look at such sequences as those listed below, we see that they can function as whole utterances or become parts of larger utterances, and thus must be considered by the native speakers of the language to be some sort of higher order units.

At home.
In school.
Come in.
Stand back.
Back up.
Is he?
He is?
He does.
They worked.
Mary Jones.
The door.

We need to deal here with two-word utterances because in terms of the learning process we saw that the children learned first to sequence two words to create utterances. Are all of those elements listed above phrases, or are some of them something else? I think that all of us who have been schooled in English grammar will think, yes, some are different. But is that really true? At this point, it is my opinion that we don't know enough about the nature of language structures to say, but so that we do not get stalled before we begin, let us at least assume that those that are complete sentences, i.e., those that have a subject and predicate, are something different from those that do not. This may not be a valid assumption, and I freely admit that it comes from my training in English grammar and not from any sound linguistic theory or practice.

No matter how one interprets the word phrase,

if we are to deal with two-word utterances, we have to admit that there are different types of such units. We did not attempt to give a comprehensive description of English minor morphemes, but what have been traditionally called prepositions do not all belong to a single class. This is easily seen when we consider an example which I once heard in a lecture given by a Dr. Bostain from the Foreign Service Institute, and that is the fact that a child of five can say to his sister, "why did you bring that book I didn't want to be read to out of up for?" Note that if one alters the sequence in which the prepositions at the end of the sentence occur, you get nonsense, e.g., read out to for up, or perhaps a different sentence. In any event, it is very clear that prepositions belong in smaller classes which have a definite sequential order when said in sentences such as the one above. Actually, what we have called prepositions in the past have very little in common with each other, except the fact that they begin a certain type of phrase, but that is not their most important function. Their most important function is that of derivation. In highly inflective languages or agglutinative languages, derivation is performed through the addition of suffixes or prefixes as a rule, but since English is an isolating language, we derive new forms by the creation of phrases. Did you ever see that in any of your grammar books? I doubt it very much, but it is true, none-the-less.

Derivation is the creation of new labels for items that the langauge does not have a word for. In Turkish, we can take any verb stem and create a noun out of it with one of several affixes. In English we create gerunds and participles in the time-honored way, but we also create something that linguists have recognized for decades and which they call composite verbs. These represent the same process of derivation, but since English is isolating, it utilizes sequences of free forms instead of a sequence of a free form and a bound form. In the list above we see the form <u>come in</u>,

which is a representative of this type of phrase or two-word utterance. Is this really a phrase, or is it something else? If we define a phrase as a sequence of words which functions as a unit, then we are forced to the conclusion that this is one type of phrase. It should, of course, come as no surprise to anyone that there would be a variety of phrases, but what is a phrase, and what do we mean by acting as a unit?

The usual answer to this question runs something like this, if a word can occupy a slot in a sentence and that word can be replaced by a sequence of words that function as a unit, then that is a phrase, because the word boundaries in a short sentence indicate some sort of syntactic boundary. To look briefly at how this works, we see below a very simple type of sentence, of the two-word variety.

Bill went.
Bill came in.

The latter sentence consists of a subject and of a predicate, the same as the former, but in the latter the predicate is a phrase which in effect derives a verb meaning roughly "to enter" from the verb-noun "come." There are many points at which linguists can argue about this composite verb, because one could argue with some validity that the verb created by the sequence is merely the meaning of the two morphemes, but this is certainly not true for many of these structures. For example, in the two following sentences, we see a totally different meaning in the derived form.

Bill came.
Bill came to.

The second sentence means to regain consciousness, whereas the first one means simply to arrive, and it is clear that the second one is a unit of some sort. Do we wish to call these phrases or not? This is not something to which I can give a clear answer for

the entire linguistic community, but for me, it has to be that these are phrases. In other words one of the basic methods of derivation in English, since it is an isolating langauge, is that of forming phrases. These phrases then are "pure verb" phrases, because they do not follow nominal distributions nor modifier distributions.

Traditionally, many people have considered the "to" following "come" to be an adverb, but we have clearly destroyed the myth that English has adverbs, but even if we had not, it is perfectly clear to anyone that this is not a sequence of verb plus a word which modifies the verb. The new phrase, or sequence, is a unit which has a meaning which is often quite independent of the meanings of either of the morphemes involved in creating the phrase. No, these are not adverbs, and any mention that these are adverbs to a student wanting to learn the language will only confuse him more than he is already by the structure of this language which does not function the way most inflective languages do and certainly not the way he probably expects it to from what he has been taught.

So, then, one type of phrase consists of a verb-noun stem plus one of the "composite verb-formers." These newly derived forms function basically as the predicates of sentences. They function in all predicate slots, even in extremely long and complicated sentences, as well as full two-word utterances.

If we stick to our definition that says phrases must always function as a unit in longer sentences, then we do not have to consider Bill went as a phrase, because this clearly does not always function as a unit. It is broken up and the meaning of the sentence is clearly the sum of the meanings of Bill and went.

There are, of course, other types of phrases consisting of two words. In the examples given earlier, we see.

At home.

This phrase is typical of a large number, such as,

In school.
Down town.
Up stream., etc.

There are, however, still other types of phrases, for example,

To go.
To run.
To eat., etc.

This could easily be called the infinitive phrase in English.

Since our purpose here is not to completely analyze English, but rather to show the nature of natural languages, we shall leave the structure of English phrases. What we were trying to show here is that there are sequences of words which act as a unit in the language. What does all of this mean in so far as the nature of natural langauges is concerned? If we look back at what we have covered so far, we see that we had a phonological level of analysis. The units on that level did not break down into a sub-phonemic level which could easily and clearly be called distinctive features which occurred in bundles simultaneously. Rather, we found a number of phonological units, called phonemes, which came in two varieties, some of which did, indeed, appear to be simultaneous bundles of distinctive features, but many of which did not. These phonemes occurred in sequences to form symbols, called morphemes, and these symbolized something for the speaker of the language. Phonemes came in two types, one generally crisp sets distinguished by distinctive features and some of which were fuzzy sets. The crisp sets set up the fuzzy sets so that they could be recognized easily through a combination of sounds and expectations of the listener. These expectations are built up by both the psychological mood of the listener and a statistical process called "Markov."

It is absolutely clear that all languages have a variety of different types of phrases. In fact, one of the most maligned methods of linguistic analysis is the "phrase-level" analysis. Linguists who utilized this method of describing have felt that when they had described the number of phrase types and the rules for making sentences out of phrases, that they had described the syntax adequately. Many others disagreed with this, because it is clear that such a description does, indeed, leave much to be desired. However, in defense of the phrase-level analysts, I would have to say that their work is probably superior to that of their critics, even with all of its flaws. In any event, the phrase system, just as the morpheme system and the word system, is a complete functioning system within itself, and since this is a fact, the linguists had a feeling of completeness when they finished this task. What we have to admit at this point is that there is a phrasal system, perhaps in all languages, but perhaps not. We really do not know, and I do not believe that we should assume this. We should explore langauge structures around the world to see if such an assumption makes any sense. At least for the languages I have worked with, something on the order of a dozen, I would say that one needs, above the word level, to go to a phrase level. Phrases, in English at least, are sequences of words (single morpheme words being counted such that both govern and government can be used as components in phrases) which function as a unit, i.e., they replace single words in longer utterances. The analysis of this level of any languages, where phrases exist, is the classification of phrases into types and the creation of a set of rules which tell us how to combine these phrases to make the next higher level of analysis, whatever that is. I am returning here to a statement made earlier that syntax is not a single level. It is a system of hierarchical levels, and we can never understand syntactic systems

by attempting to analyze them as single systems, because we will be attempting to handle elements from a number of divergent systems with different basic units and different rules for combination into meaningful sequences.

To look just briefly at the phrase level of English, we see at least the following types of phrases.

Single word phrases (any English word can act as a phrase on this level. Without this provision, there is no hope of understanding this system in English and probably in all languages).
Infinitive phrases, e.g., to run.
Prepositional phrases, e.g. to the house.
Composite verbs, e.g., come to.

I do not at all pretend to understand how the English system of phrases, i.e., the phrase level, works, because I have not really gotten that deeply into it, but one can see from the above list that the actual phrase types in English are somewhat different from what you might have heard. We have the complication, for English, that being an isolating language much derivation is replaced by sequences of words rather than sequences of morphemes, remembering, of course, that morphemes are also words in isolating langauges, and that is precisely the reason that these are not suffixes. I cited the above list only to show one the nature of phrases and a few examples of various types of phrases. Note the absence of a "nominal phrase" type. There is, perhaps, a noun-verb phrase, but certainly not a noun phrase.

Now, we have seen that the basic kernel sentence type in English is the single word and that all morphemes, except a very short list of suffixes, can function as whole utterances creating sentences of this type. Many whole utterances are also created by words, i.e., sequences of a stem plus an affix, but also many utterances consist of two words, often a major morpheme and a minor morpheme or a word, which consists of a stem plus suffix, and a minor

morpheme. Some phrase types are very easy to define and understand because they begin with a specific type of minor morpheme, such as the prepositions, and end with a specific type of morpheme. This makes analysis easy, but some are not so easily dealt with, so we leave the phrase level of analysis and move up to three. word utterances.

Three-word utterances are exemplified in the list below for English.

To the store.
I am going.
Bill eats garlic.
Mary is short.
The big one.
Big and ugly.
Big black balls.
The rolling stones.
The burned buildings.

Above is probably not a complete list of English three-word sentence or utterance types, but they will serve as a springboard for us to consider the nature of three-word utterances.

Clearly, the first example is simply a prepositional phrase, which could be utilized as a complete utterance in response to the question, where are you going? Such phrases begin with an element from one of the paradigms of composite verb formers and end with a noun-verb, in between there can be a long list of modifiers, but there is usually not more than four or five. Concern by some linguists that one needs a system for creating infinite sequences of such is not worth noting, because when the list passes five or six, most listeners lose interest and do not follow the conversation, so for all practical purposes six or seven is as many as we can put in.

The second utterance brings us to a definition of a sentence. A sentence differs from other two- or three-word utterances in that it has a true pre-

dicate. A predicate is a word which is inflected to show past time and 3rd person subject or has a special irregular form to do that. Note here that this precludes such forms as going, in Bill is going from being the predicate, which I believe reflects clearly the intuition of the native English speaker before he has been taught a bunch of garbage in an English class. Also, we note too that the so-called past-tense forms of regular noun-verbs and irregular noun-verbs, in sentences, such as, they were burned, are clearly not predicates. English is a funny language, from the point of view of many foreign learners, because we tell them that the form they were burned is a passive voice and that burned is the predicate, yet their intuition about languages tells them that this cannot be so, because if it were indeed the predicate, burned would be inflectable for person and tense, which it clearly is not. This brings us back to the analysis of the word-level in English. If you have followed our line of reasoning up to this point, you will remember that we said that English has two word-classes which do not parallel any morpheme-class, but rather, perform syntactic functions of both noun-verbs and modifers, but not all noun-verb functions, and one function which neither noun-verbs or modifiers can perform. What we see then is a language which does not have anything comparable with a passive voice at all, or at least it is absolutely clear that forms such as they were burned are not passive verbal structures. These are exactly parallel with, they were tall. This is easily and clearly demonstrated by comparison with such sentences as, the burned men died and the tall men died. Note that the words tall and burned are performing exactly the same type of semantic and syntactic function in both types of sentences. Since we have a class noun-verb which performs both nominal and verbal functions, in traditional terminology, how does a listener know what is happening. We need a set of rules, then, to tell us what is the predicate

in sequences which include a number of items which could be the predicate. Such rules are easily formulated for English, but as an example, they would go like this: gerunds are never the predicates of sentences, forms of to be are always the predicate, except for the infinitive form, and so on. Rules about the forms of to be always being the predicate are necessary, because burned and other participial forms are sometimes the predicate, but not in the example given above, because the form of to be is functioning as the predicate, and it is absolutely imperative if you are to get the correct meaning out of a sentence to know that the participial forms are or are not functioning as predicates in English. Relative to three-word utterances, note the vast difference between; it was burned, and he burned it. In the former case burned is performing a modifier function and the past time is shown by the form was. This could just as well be; it will be burned, which is clearly not past time, but foreign students will think that it has to be past tense if they think that burned is the predicate, because participial forms do really mean past time if they are functioning as the predicates of sentences. The second example above, i.e., he burned it; is an example of that. One of the main functions of the articles in English is to inform the speaker that the word following them is not the predicate of the sentence. We often say that this signals that the following form is a nominal, but that is not entirely true, because what is a nominal? Clearly, however, it signals that the following form cannot be the predicate, and this is one of the most important things for a listener to be able to recognize when he or she hears a symbol which is a noun-verb.

Returning now to our list of three-word sequences, we see that some have predicates, and these we can label sentences, but the others could be called different types of phrases. In sentences such as; they were destroyed; the subject they can be re-

placed by the phrase the burning buildings or the burned buildings, as well as a large number of other morphemes, words and phrases. If we make a substitution such as this we get the sentence, the burned buildings were destroyed.

Now, this raises the question as to whether the human brain thinks in some sort of deep structure, which is independent of the language it is using, that is, something that is a natural part of the human biology, or does one merely feel and see images in terms of his sensory mechanism and then convert these sensory responses into the symbols that the surface structure of his langauge provides for him. If the former is the case, then what a person does when he thinks is see or feel something, for example, a picture of something resting on a cook-stove cooking. Then he formulates what he wants to say about it directly in the surface structure of the language which he wishes to speak. This logically implies that all complex thoughts or thoughts about things which he cannot sense immediatly are created directly in the surface structure of the language he is going to speak. On the other hand, if there is a "language" of sorts built into the human mind, then the transformationalists are correct, and he thinks in that language and converts that into the surface structure of the language he is speaking. This is the curx of one of the greatest arguments in the field of linguistics over the past two hundred years. The entire concept that a person has a language system wired into his brain, Chomsky's deep structure, is just too bizarre for me to accept. If such a complicated wiring actually existed, how, in God's name, could all of the diversity in the surface structure of languages which is found around the world exist? Certainly, if there were a deep structure, similar to that postulated by Chomsky, there would be a great deal more similarity in the nature of langages observable around the world than we find. No! It seems absolutely clear to me that the only similarities which we actually see in langauges are

those which result from misanalysis and distortion left us as a legacy of the Latin grammar with which the missionaries and linguists of the nineteenth century went about the world describing the languages. Much of what we have been saying about the nature of English, and our reasons for saying it to this point is to show just how erroneous such a picture of this language actually is. There is something which we call language structure, that is, a grammatical system, which exists within the language itself. The function of the linguist is to find exactly how that system works in each language and, hopefully, from a vast amount of knowledge of this sort, we can in the distant future, learn something about the similarities of language structures. Now we have such absurd statements of similarities as, all languages have nouns, or all languages have this or that other thing, and we abslutely do not know this. In fact, the evidence from English, Turkish and Japanese alone, leaves little doubt that this is false, because what we call nouns in Turkish includes all of what we call modifiers in English, and the class noun in English includes most of what Turks would call verbs, so where is the validity in calling such diverse things the same? About all I can say about the Japanese system at this time is that talking about it in terms of nouns and verbs leaves one so confused that he would never understand anything at all, but I do not know what the real classification is. How, then, does the human brain "really" work, if it does not work in terms of Chomsky's transformations from deep structure to surface structure? We shall leave that for the last chapter of this book, but for now, let us just say that it seems very very clear to many of us, from the "nature of natural languages we have studied around the world" that what language universals actually exist are very different from those which have been published as existing and that there is no "deep structure" as postulated by the transformationalists. Whatever deep structure exists is a

much more flexible mechanism. This must be true because if everyone had a deep structure which is as rigid as Chomsky proposes, then there would be no reason for the various speech communities around the world to have modified this basic mechanism. All languages would just be like that. There would be no need for transformational analysis, because we would have simply a large number of languages which were nearly identical in structure, perhaps utilizing different words and different sounds which resulted from the sound-shifts of the ages. Clearly this picture bears no similarity to that in the real world of languages. The deep structure would appear to consist of two sets of storage, such that whatever sensory experiences are stored in one, a parallel set of verbal symbols can be stored in the other one in such a way that the two can be continuously associated for a very very long period of time. Further, there must be some sort of self-programming of the search pattern, but the nature of the search pattern is completely independent of anything else and is built up by the speaker of the language as he learns the language. It is not wired in. There appears to be no limit to the different types of classes in languages, only the basic point that all languages do classify their morphemes words, phrases, etc., into types or classes. This is true and necessary, because only through having such classifications can human language work, but each langauge has a unique set of classes of all items and the basic wiring of the human mind does not place any restrictions on this classification at all. At least none has been demonstrated even slightly to this point. Virtually all of what we note as such similarities in our linguistic literature is the result of the history of the field of study, not the nature of real languages. It is necessary now for us to reanalyze languages in terms of their "real" classes. Then, and only then, will we be able to discover what is actually universal. Then, too, we will have a much better picture of the nature of the deep structure.

Let us return now to a discussion of the nature of utterances longer than single morphemes or single words. We have been looking at the way languages are learned and the way longer sentences are built up in terms of one-word utterances, two-word utterances, and so on. It seems highly probable that if there is any basic wiring in the brain that that wiring will make itself evident in the way langauges are learned, and since all recent studies indicate that children do appear to build up their ability to utilize a language in this way, whatever basic wiring exists should be found there. One-word utterances, in English at least, appear to consist of all morphemes in English, except for about fifteen or twenty affixes. Two-word utterances are somewhat more restricted. That is, not every morpheme can occur in second position. Such restrictions must be stated in terms of classes, not individual morphemes. For example, I heard a long lecture by a transformational grammarian in Singapore one day trying to explain the reason that we could say the dead man and the burned building, but not the killed man. This sort of discussion is futile and pointless in the extreme, because one certainly can say the killed man. We do not usually do so, because we do have the modifier dead which fills the bill very nicely. However, if there were two or three bodies laying in a row, and all of them had died of smallpox, except one, I would say, "bring the killed man to the morgue for an autopsy." This is perfectly clear English. It is not bad grammar, bad style, or bad anything else. All we are dealing with here is the statistical utilization of words with very similar meaning. In most situations, the dead man would be applied to the killed man. The only exception would be where one needed to make a distinction between one who had been killed and one who had died of natural causes. Most of us have little reason to do this in everyday life. Hence, the killed man is a very unusual sequence. The point that must be made is that inside the brain, linguistic units are stored

in terms of their classification, and when the brain needs a word, morpheme, phrase, or whatever, it goes to the spot wherein those elements that can occur next in sequence are stored. Then it goes to the visual image or other sensory experience it wishes to express and finds the one that is most appropriate. Often this is a rather broad label. The label will be only minimally specific, i.e., specific to the point that the listener can determine what is meant through a combination of sensory experiences, linguistic knowledge and his own past history, i.e., the shared experiences of the speaker and the listener. The relevance of all the above discussion to what we are considering at the moment is that many morphemes, words, as well as, sequences of words and morphemes, never occur simply because the speaker never had any occasion to express a certain thing. This must not be confused with grammar. Grammar is a statement of the classes of elements in a linguistic system and the rules which determine how these elements are put together to form meaningful utterances. The fact that one does not often say, I hit me, does in no way mean that this is ungrammatical nor even that it is not good style. I hit myself is considered to be better style by most educated English speakers, but that does not make the former sentence, i.e., I hit me, ungrammatical, and any grammar which attempts to write its rules in such a way as to eliminate this latter sentence is not only going absolutely to fail but if it were so written it would be a gross distortion of the nature of English. Statistical probability of usage is not a determinant of linguistic analysis, though statistics are sometimes important in some facets of analysis.

Our method of analysis then, in a way, is a reflection of the way children learn their first language, since this is probably based on the natural wiring in the brain. Our method of analysis then must become the sorting of forms

into classes of units, i.e., words, morphemes, phrases, whatever, which mutually replace each other in sentences. This means that notions about whether a noun is a label for something alive or not, and hence is an animate noun, applied to the analysis or description of a langauge is both useless and misleading. Here again we have the situation that English does not make any such distinction, Algonkin languages do, but many of the things in the class "animate nouns" in Algonkin languages are not "alive," e.g., glass in some languages is animate. Statements such as, every human being can tell the difference between animate and animate things in the world, hence all languages make this distinction are absolutely wrong and any linguistic theory which is based on such "meaning" or "reality" distinctions is doomed to absolute failure, because languages do not work on any such basis.

At this point, let us look at a few more three-word utterances, because we have seen that some three-word utterances are phrases, some are sentences and some could be considered to be only sequences. For example:

just sequences;	the big house, the yellow dog;
sentence;	he will be burned tomorrow;
phrases;	up the street, to go, come to;

Look now, also at:

While running home.
Running to school.
Engulfed in flames.

Note the last three sentences are different from any of those given in the first list. How are they different? They do not have a predicate, so they cannot be sentences. Remember, we established the fact that gerunds and participles could not be predicates of a sentence earlier in the chapter. However, we feel intuitively that here we have something like a predicate

forms, and, as a result, I think that most native speakers of English would not like to include these in with either phrases or sentences, because they feel that it is a different kind of unit. It has in it a predicate-like word, which is clearly not a normal predicate, since it cannot take any inflections at all, but still seems to do for this type of unit what a predicate does for a sentence. What then are these units to be called? We might just as well call them clauses as anything else. That is what they have been called traditionally. A clause then is a sequence of phrases and words which function as a unit but which is an element within a sentence. Here we begin to see a fairly clear picture of the way English, as an example of languages generally, functions. We see morphemes combining to form words, e.g., go + ing, to form going, then words form phrases, e.g. to the store forms a phrase, phrases form clauses, e.g., riding + up the street + in his car + to school forms the clause riding up the street in his car, and, finally, sequences of clauses form sentences, e.g., riding up the street in his car to school + looking for his sister in the street + was indeed very exciting. Utterances, then, are made up of units on several levels, or so it would seem. A sentence can consist of a clause, such as, while going south to Mexico, plus a morpheme, such as, he, plus a phrase, such as, will be, plus a clause, such as, driving his old oldsmobile, plus a phrase, such as, in second gear, plus a morpheme, such as, most, plus another phrase, such as, of the time. This would, of course, form the utterance, While going south to Mexico, he will be driving his old oldsmobile in second gear most of the time. This would then be a sentence which consists of an introductory clause followed by a single morpheme which is the subject of the sentence followed by a complex phrase, which would have to be a type of phrase as we have stated the nature of languages, though it is not exactly what one might consider a verb phrase in traditional grammatical terminology. Then

we have a complement, driving his old oldsmobile most of the time. This shows the hierarchical nature of languages. Clearly, there are a number of systems in languages, then. From the bottom up, we see phonemes, then morphemes, then words, then phrases, then clauses and finally sentences, and no one really knows what levels may exist above that. While many of these elements have been recognized for decades, most people still attempt to handle all of the language structure above that of the word as if it were a single self-contained system. Further, much confusion has resulted in linguistic theory because traditionally we have defined morphology as "word analysis" and we have seen absolutely that if one does not look carefully at a morphemic system and a word system as two separate systems, he will obscure some very very important aspects of the structure of the language. Then above that of the morpheme and word, we see that many linguists have treated all syntax as if it were a phrase level only, and some other linguists have treated it all as if all you need to deal with are sentences, and this too obscures some very important aspects of the linguistic structure.

If we take a brief look historically, I believe that we can see the history of linguistic development, and I also believe that this will tell us some very important things about the way that we absolutely must go if we are to understand how languages function and learn to utilize them effectively. First, the Greeks, and their students the Romans, considered everything to do with language to be a single mess of data. They studied all aspects of linguistic structure and also included logic. It was not until about 1850 that the field of linguistics was able to break out of this and recognize that it was essential to look at language structures in terms of at least two separate systems, that of phonology and grammar, the latter including everything dealing with the meaningful elements in language. Then by the beginning of the twentieth century we see the break up of

these two "levels" into "phonetics and phonemics" and "morphology and syntax." It literally took 2,000 years of study to force us to realize that language was not a single system but a set of layered systems. Then in 1936 George Trager defined phonetics out of linguistics with the general statement that it represented unstructured data, hence it could not be a part of a linguistic system. Then Jakobson et al reintroduced a sub-phonemic level in the form of the distinctive feature level of analysis in 1951. However, we can see now that the phonemic system does not easily break down into such a two-level system, but rather consists of two types of elements one of which appear to be traditional crisp sets, distinguished by distinctive features and some of which are fuzzy sets. These seem to coincide rather well with the traditional breakdown into consonants and vowels. Then on the level of meaningful units, we see that there is a clearly separate level of morphemes and words, because languages have different classes of items on the morpheme and word levels, and throwing all of these into one ball-of-wax is a very good way to get very confused about the way the system or systems work. Above the word level we see that there are additional levels, none of which have been studied properly as yet, and this is for the future.

Concluding, then, we see that as we have learned more and more about how language works, we have found that it consists not only of a single system but of a number of more or less independent systems all of which must work together and in harmony if the speaker is to communicate with his listener. Not only are there a number of systems, but these systems have a hierarchical order, which is also a necessary part of the nature of natural languages. What we need in the future, then, is a careful study of exactly how many such levels exist and how they fit together to form the system which we call language. In the next chapter we shall discuss the way all of these things work when we talk to each other.

CHAPTER VI

So far we have seen something about how the sound systems of language work, i.e.; how phonemes are differentiated from each other by distinctive features and as fuzzy sets; how morphemes come in two fundamentally different types, some of which form crisp sets and some of which are quite fuzzy; how morphemes are combined into words which are not necessarily organized into classes in the same way that morphemes are; and finally that on the so-called syntactic level of analysis there are a variety of different types of units, most of which are very poorly understood at the present time. Now we need to see how this all works together to make communication possible.

First of all, it is imperative, if you are to understand language systems and how they work to accept the basic fact that in conversation there is no communication of information. This is extremely difficult to get people to accept, because they are used to seeing definitions of language most of which have communication in the definition. What is transmitted from person to person is sound and sound alone, so far as language is concerned. However, language does not operate in a vacuum. Languages are always utilized in a situation, and that situation adds enormously to the perceptions of the person who is listening in a conversation. Also, each person brings to the conversation a life-history, and this life-history adds to his interpretation of many things. We have to realize that I make a noise, that noise moves through the air and is picked up by the ear of a listener. Then the fun begins. The listener begins to process this noise to see if he can possibly figure out what it was that I wanted him to know. It is

this process which we propose to explore at this time.

Beginning at the beginning, where everyone should always begin, we start with the movement of the vocal chords of the speaker. This stirs up the air in a systematic way. This systematically highly organized wave motion goes across the room and impinges on someone else's ear. The listener must process the data from "left to right" as the saying goes, because that is the way it comes in. There is absolutely no possible way that he can do otherwise. We can clearly demonstrate that this is the case, because if we stop a sentence in the middle and ask a listener what has been said up to that point, he can always repeat all of the words that he has heard. Since we can stop at any point and he can always identify what has been said up to that point, this logically implies that he has clearly identified each morpheme in the sequence as it came in or very very shortly thereafter. For example, I can stop a person who is talking after each word (not morpheme) as in the following sequences.

I.
I am.
I am going.
I am going to...
I am going to go.
I am going to go in.
I am going to go in a...
I am going to go in a minute.
I am going to go in a minute to...
I am going to go in a minute to the...
I am going to go in a minute to the store.
etc.

Since people can do what has been illustrated above, it is absolutely impossible that they could do this unless they process each sound and each morpheme as it comes into the brain. Hence, let us discard the myth that people do not process linguistic data from "left to right," by which the advocates of that idea mean in a sequential order following the way it comes

into the ear.

Before we proceed further, it is imperative to make the point that since the only thing that is transmitted from speaker to listener is the vocal symbol, i.e., the form, all linguistic analysis must procede from the forms. By this we mean that one must not begin the analysis by looking for a first, second and third person pronoun singular and plural and then hope that he can straighten out the mess later on. What has to be done is to find the forms that occur in that language, as we did earlier for English pronouns. We find the forms, I, you, he, she, it, we, they. If we begin with this set of forms and attempt to understand how these operate in the system we will learn much more quickly how that aspect of the morphological system works. Actually, if you remember, we discovered that English marks the speaker as the sole subject, I, the speaker as a member of a group which is the subject, we, the listener whether a member of a group or not as the subject, you, then it is imperative in this language to distinguish groups from individuals who are not the speaker or listener, plural, they, singular, masculine, he, feminine, she, and finally, neuter, it. It seems almost certain to me that this is the way the human brain distinguishes person in English when he speaks and hears the language. It is a distortion of reality and confusing to students to say that we have a you which is singular and you which is plural. Then the point of departure for analysing any language is the sets of forms which exist in that language and not some preconcieved system which it would be nice if that language followed. Remember that our purpose in analyzing a language is to attempt to reconstruct what goes on in the mind of the members of a given speech community as they attempt to communicate through the use of this complicated symbolic system.

Somewhere in the brain of the listener it appears that he has stored a statistical table which tells him which sounds can occur in initial position and which sounds cannot. This makes it pos-

sible for him to much more quickly identify the incoming phonological unit because the brain does not have to attempt to compare the incoming signal with every possible phoneme in the language but only those which can occur in initial position. Further, there is a considerable body of evidence to indicate that the brain probably has some built-in method of comparing those most frequently utilized sounds first, making succeeding comparisons in order of descending frequency of usage. This would, indeed, reduce the time required to identify a sequence of sounds enormously.

How is the above suggestion possible? Actually, it is probably something quite simple. It is quite possible that each time a unit, on any level, is used by the speaker his mind has to touch with some sort of electrical charge the storage compartment in which that unit is registered. When this is done, it could add a little more of whatever chemical or electrical energy is utilized to make the charge possible in the first place. Then all the brain has to do is have some mechanism of causing itself to look at those elements with the highest charge first. It could be that while you are asleep at night, the brain reorganizes the storage in a sequence with the highest charges stored at one end and the weakest charges at the other end. Forgetting may simply be the fact that some items become so weak that they are not shifted in this shuffling, because the charge weakens when it is not intensified by occasional use. This might well account for the fact that many people seem suddenly to be able to understand things in a foreign language that they could not a few days before, and without any additional study or effort on their part.

If that initial sound is a consonant, then the brain simply looks for the presence or absence of distinctive features. As soon as it has identified the presence of a fixed constellation, then it can confirm that the initial sound is a certain consonant. However, if the initial sound is a vowel,

the brain must compare the input, in terms of frequency concentrations in the spectrum, with norms for each of the vowel phonemes in that particular language. I believe that the brain identifies the correct phoneme by something that must parallel a distance computation in a three dimensional space. If you remember from your high school algebra, distances in any three dimensional space can be computed by subtracting the location (in this case the cycles per second of the center of Formant I, Formant II and Formant III on spectrograms) of a point in each of the three dimensions from another such point. Then by squaring each of these differences, adding them together and taking the square root, we get the distance. I do not for a moment think that the brain does this, but I believe that it does something that is very much like this. Once the brain has a measurement similar to the one just described, it determines which of the norms is the closest and identifies that phoneme. In English, since all phonemes, except /ʊ/, as in book, and ng as in sing, occur in initial position, the brain has to make a sequence of compares begining with the phoneme utilized most frequently in initial position. Such compares in a computer would take about one one-thousandth of a second. Since the average phoneme is more than five one-hundredths of a second, the brain, working at this speed, and it probably works much faster than that, could make at least fifty compares while a single phoneme is being uttered by the speaker. Hence with s very short delay, one phoneme could be identified while the next phoneme is coming in. English has about 43 phonemes, so the brain could easily make a compare for each sound and still identify the phoneme before the next sound is completed but since it begins with the ones utilized most frequently, it would only be a very rare occasion on which it would have to make the full 43 compares before identification were completed.

Once this initial phoneme has been identified,

the problem then is to identify the second phoneme. This could, of course, follow the same procedure as for the first, but much evidence appears to indicate that this is not true. As we asserted earlier, every known language has restrictions on the distribution of phonemes. This is probably there because of the basic wiring of the brain, that is, the way in which the brain handles data. Taking English as an example again, we see that if the first phoneme is identified as a stop consonant, then the second phoneme in a word or utterance cannot be a stop consonant. This means that the brain would begin at the top, that is, with the sound that has the greatest frequency in second position. If, after having gone through the entire inventory of sounds that do occur in second position, it does not find a match, then it will try the voiced stops, if it has clearly identified that this sound is a crisp set (consonant like sound) as opposed to a fuzzy set (vowel-like sound). This could easily be done simply by noting whether the incoming sound consists of distinctive features, such as stoppage, friction, etc., as opposed to being an array of concentrations of sound at certain frequencies in the spectrum below about 2,000 cps. For those who are interested, cps means cycles per second, and there is a technical term for this. It is called the Hertz, but I prefer to continue to use cps, because the abbreviation will help those not familiar with acoustic theory.

To give a clear example of exactly how such an identification could be made, even by a man-made computer, let us look briefly at the spectrum of a sibilant. The sounds /s, z, sh, and zh/ as in the English words; sip, zip, ship and leige, are all characterized by a wide and very strong band of more or less randomly distributed noise above about 3,000 cps. If one of the two voiced sounds, that is, /z/ and /zh/ is produced then there is a strong concentration in the lower frequencies, if /s/ and /sh/ is produced then there is not. The difference between the /z/ and /s/ and the so-called "blade sounds," that is, /sh/ and /zh/ produced at the same point of articulation is that

the former have a much stronger concentration of noise in the upper range of the spectrum whereas the latter have a stronger concentration between about 3,000 and 4,000 cps. All one has to do is to subtract the volume, as measured by the strength of the incoming signal, of those frequences above about 4,500 cps from those below it. The computer can then tell by whether the sign is plus or minus, depending on which way you subtracted, whether the sound is one of the two former sounds or one of the two latter sounds. I do not wish to belabor this point here, because I suspect that most readers are not particularly interested in acoustics. I just wanted to point out that many of these types of identifications could be made very quickly and easily even by machines. Let us return now to the sequence of identification.

We have defended the much-maligned Markov process because we believe that it is absolutely impossible for languages to work without it. As in the second position, once the brain has identified the second sound on the basis of a combination of the the incoming sound pattern and the statistical probability that a given sound can or cannot occur in second position, then it must face the problem of the identification of the third sound. To show exactly how the Markov process is important to the identification of phonemes, remember that we said earlier that the English vowel phoneme in book, which we symbolized with /ʊ/, did not occur in initial position in English utterances. If the listener's brain identifies the sound that comes in as closest to that vowel phoneme in its process of identification, it will assign the vowel to the nearest "other" vowel, because it knows ahead of time that this vowel simple cannot BE in that position. I see absolutely no other way to account for the types of acoustic variation that I have seen in my laboratory combined with the reports from the Haskins laboratory and a fact reported to me by Martin Joos once in a class I took with him on acoustic phonetic. That fact is that one could filter out enormous quantities of English phonemes

in normal speech so long as the initial phoneme and all vowels under primary stress in English were not filtered out. These researches were reported to have been carried out at the Bell Telephone Laboratories, though I have not read those reports. Here again, we see strong evidence that the search and identify process is strongly involved with the way in which English phonemes are sequenced, and while there is no evidence that I have seen to show that one or another type of sequence is wired into the brain, we do see that the fact of having a limited number of possible sequences, clearly and sharply defined in the nature of the language, is absolutely imperative to communication, and this clearly implies that the Markov process is essential to any sound theory of language or communication.

Now, to further clearify the way the Markov process works, we must remember that when we state that we have, in a given position, a given probability that a certain sound will occur, this means something very different from several examples that I have seen in the literature which purported to exemplify the process. If the initial phoneme is a voiceless stop, then the second phoneme in a word cannot be a stop of any kind. Once the brain has identified a sequence of /sp/, for example, then it can eliminate virtually all consonants except /l/ and /r/, so that the search has only to cover those two phonemes and the vowels. However, not only are most of the consonants eliminated following this sequence, but so are some of the vowels. The process of vowel identification will be the same as given above. The phoneme will be identified as the sound with the norm nearest to the pattern of incoming sound. If that sound is an impossible sound in that position, then it will accept the next closest "other" sound. If two sounds are equally distant from that point, the brain will simply hold until it gets some further information. What must be remembered is that once one phoneme has been identified, each following sound has different possibilities of occurrance, once the second phoneme has been identified, each

phoneme has different possibilities of occurrence depending on the specific sequence which precedes it, e.g., certain sounds are impossible or extremely improbably following /sp/ and quite a different arrangement of sounds is possible or probable after /pl/. I contend that languages would be extremely inefficient, if not impossible, without this type of arrangement, and it is precisely this type of system which accounts for a great deal of free variation, particularly in those sounds which are identified on the basis of being members of fuzzy sets, such as, the vowels.

Suffice it to say at this point that this process continues until a morpheme has been identified. In English usually one can identify a given morpheme after about four or five phonemes have been identified, depending on how much additional information is supplied to the listener by the environment, etc., regardless of how long the word or morpheme really is. When a particular morpheme has been identified, the brain resets, not to zero but, rather, to a new set of probabilities for sounds which initiate morphemes, but not necessarily utterances.

Now, it is necessary to exaplain such utterances as /mgoin/, which I used to say each morning as I stood in the front door, hat in one hand and briefcase in the other. This, of course, meant, "I am going." We note immediately that there is a /g/ following an utterance initial consonant, /m/. How does this fit into our theory of how language works? Clearly, if the brain searches through the possible sounds in second position and finds none, then it knows that it is dealing with the end of one word, which it may simply not have heard or which may not have been said, so the brain makes the assumption that it did not hear. Therefore, if none of the possible second position sounds can be clearly identified, then it must search through those that are not probable or possible in that position, so it will then go to the voiced stops. The instant it identifies /g/, it must hold its identification until it can tell what word was in first po-

sition. What are the possibilities preceding an /m/ phoneme? Very few, indeed, e.g., while going, when going, is going, etc. Note that of those possible forms listed, none end with an /m/ phoneme. The only possible word which has any sound near the /m/, is am. Again, in terms of the statistical probabilities preceding am, that form of to be can only follow I. Hence, within certain limits, the Markov porcess works both forward and back. Further, if the brain comes up with two sequences of morphemes which appear to be the same, it must then look to some other sensory input for the answer. What we are saying here is that the incoming signal is a sound signal. The analysis and discussion of language must proceed from the symbols used in that language. When the listener hears an incoming signal, he interprets the meaning of the sequence of symbols in terms not only of the incoming sounds, but of a statistical table of sorts which has been built up inside the brain somewhere and if it fails to make a unique identification on the basis of sound, it falls back on the statistical table, and if that too doesn't work, then the brain must fall back on what is probable in a given environment. The fallacy in much linguistic research in the past has been to attribute far too much of the meaning content in a given sentence to the forms alone, when it is perfectly clear that much of the meaning content comes from things independent of the sounds themselves.

Rather than bore the reader with all of the details of each step along the way toward interpretation, let us suffice it to say at this point that phonemes are identified, more or less as described above, then morphemes are identified in terms of the sequence of phoneme combined with the statistical probabilities of occurrance. Then morphemes are identified on the basis of the phonemes and a Markov process (statistical probabilities) until a word is identified. Then the phrases are identified on the basis of the words and a statistical process and so on up the line.

All of the aspects of communication are involved on all levels, and these include language, paralanguage, kinesics, life-histories, social context, physical context as well as Markov probability. It is this combination of factors which makes it so difficult for people to sort out and discover exactly what is doing what.

Once a given morpheme has been identified, then only certain other morphemes are possible in a given situation, others are considerably more probable than others, and all things being equal, the more probably items are likely to be "heard" than those which are improbably in a given situation. For this reason, more careful articulation is necessary, and produced, when a person is giving two words like "latter" and "ladder" in a situation wherein both are likely, but we do not articulate so carefully when only one is possible. Much of linguistic interpretation of symbols comes not from the signal itself, but from the situation in which the word occurs.

Since the meaning is often not carried by the language or "deep structure," it is necessary now to look at some examples which I have garnered from the literature in various places to see just how bad linguistic analysis can be when people forget that laguages operate in a real situation. Let us begin with one of the Transformationalists favorite problems, that of solving ambiguities. One such example which I have seen repeatedly is:

The snow was piled by the students.

It is said that there are two deep structures, and that it is useful to people studying language to know that the two meanings, (1) the students did the piling and (2) the snow was piled beside the students, are derived from different deep structures. I fail to see how! If I am listening to a person who tells me the above sentence, there is one and only one way that I can tell which meaning is intended, and that is to see the situation. I

cannot possibly tell from anything inside my head, be it deep structure or anything else, nor from anything in the structure of the language itself. I must look out and see the students with shovels in their hands or a pile of snow beside them. It serves no useful purpose, so far as I can see, to suggest that one form is a transformation from the active form, the students piled the snow, and becomes a passive form of that structure. Clearly, this is not a passive formation of anything at all. The sentence is best broken down in this manner:

The snow + was + piled + by the students.

The first unit is the subject, the form was is the predicate, the word piled is a participle which modifies snow (note, the piled snow was by the students), and this in no way constitutes a passive form of anything. Finally, there is the phrase which simply adds further information about the snow. This information is a phrase which tells us about whatever went before, but we cannot know in what manner, except through knowing the situation or sometimes directly from the meanings of the words themselves. The critical question above is, does "by the students" modify the predicate, was, or the complement, piled, and we can tell this only from extra linguistic information, so postulating a deep structure helps us none at all. Now let us look briefly at a few items wherein one can tell only from the meanings of the words themselves, and here again, deep structure is something that we need like we need a hole in the head. Note the following examples taken from the *Foundations of Linguistics*, op. cit., p 153.

This actor reads well.
This poet reads well.
This book reads well.

It is said that IC analysis is inadequate because it assigns only a single type of constituent structure

to all three of the examples given above. I say "hurray for IC analysis," because that is exactly the way the structure works, and any kind of analysis that yields anything else is distorting the structure of the language again, just as we were earlier when we were discussing other facets of grammar. One knows that the actor reads well because one does not read an actor, unless he has tatooing all over him, in which case the above sentence would be perfectably acceptable, and it is precisely this fact which mitigates against any statement that these are different types of sentences. They are identical, and the only reason that we resist this is that we have been brainwashed by years of English in our school systems. This brainwashing makes it impossible for us to see the necessity for seeing these sentences as exactly alike. This relates to what I said earlier when I stated that many people say that the structuralist approach is inadequate, basically because it comes up with something that disagrees with what we have been taught all our lives. The three examples given above are absolutely identical, and the differences in meaning come absolutely from the meanings inherent in the words which comprise the sentences. One interprets the overall meaning of the sentence in terms of his associations with those words, and if he had just seen the movie "the illustrated man" then "the actor reads well" might very well be interpreted that someone was reading the actor, and that in no way alters the nature of that syntactic structure.

I am not intent here on defending IC analysis, but rather I would like to see if we cannot understand how languages work. In the work cited just above we also find two sentences:

John killed Mary.
Mary was killed by John.

It is said IC analysis fails to show the relationship

between them, and perhaps one way of looking at it would, but when the authors say that one needs to show the relationship between "active (5) and passive (6) sentences," I would disagree radically, since we have seen that there is no passive in English. This being the case, we certainly do not want a grammar which shows such a relationship. The way English really works is that in the so-called passive sentence Mary is the subject, was is the predicate and killed is the complement, which simply describes Mary's condition. Then there is the added phrase by John which tells us who did the killing or that she was killed at his side, and once more we cannot tell this from anything in the linguistic symbols themselves. We must know this from previous experience or be able to see the two bodies lying side by side. The error made by IC analysis is to consider killed as the predicate, that is, the head of a verbal phrase, when clearly it is not. This error is a result of listening too carefully to traditional English grammar.

I do not wish to pursue this question too much here, because we are interested in how languages work, not whether IC analysis of English is correct or not. What I wish to show in the examples above is that much of the ambiguity, which presumably stimulated much research into transformational grammar, can be resolved only in terms of the situation and only if it can be experienced by the listener or from the meanings of the words themselves, and I think that the illustrations above do that adequately.

Then we see on all levels of analysis the simultaneous operation of all of the elements listed above, i.e., phonology, morphology, words, phrases, clauses, sentences, combined with life-histories, the physical and social environment and the Markov process. Any theory about how languages work will be quite inadequate unless it attempts to pinpoint exactly which element of meaning is coming from exactly which of the contributing factors listed

above. To show how this must be done, consider the following sentence.

I hope Mary will go home with Bach.

The first thing that happens is that the incoming sound is processed and the pronoun I is recognized. Following this pronoun the listener knows that if a member of the morpheme class noun-verb in English occurs, it must be functioning as the predicate. He also knows that are, am, is and were along with hundreds, perhaps thousands, of other morphemes cannot follow I. Hence the search pattern is narrowed. Once the I has been recognized as an independent morpheme, because there are no words in English wherein an /h/ follows an /ay/ and hence there has to be a word boundary after I, the sounds of hope are processed sequentially. After /h/ there are a limited number of possible phonemes, hence the vowel /o/ can be identified rather quickly, and since /p/ is a bundle of distinctive features, it too is easily recognized. Now, we come to a fact about English that is rarely mentioned, and that is that certain words, such as hope, think, wish, and many others reset all of the Markov indicators back to the same as they would be if you were beginning a new utterance, because following these words, one can imbed a sentence, i.e., we can absolutely begin the part of the sentence following these morphemes as if we were beginning a new utterance. Then Mary is processed and recognized as the subject. Now we come to a modal, and most books will tell you that this is a future tense. But? Does will mean future time? The answer to this is, absolutely not! This statement is often made while Mary is on her way home at that instant. Further, when we say, I'll go, we often are in the process of leaving as we make the utterance, and this means that all will indicates is that the time is non-past. Actually, the modals express something quite different from what we have been taught. Each modal tells something about the degree of certainty of an action. These forms have absolutely nothing

to do with time. The only reason will has ever been associated with the future in English is that it implies intent, and one rarely intends to do things in the past. However, that it is independent of time can be shown by literally thousands of examples besides the one given above. Another common one is that a mother comes to the side of her child and begins tying her shoes, which the child has been failing at miserably, and as she is tying them, she says, I will do that for you. A young Viet Namese student, Bui Thi Hong Nguyet, did an excellent study of modals while in Singapore in the SEAMEO RELC in 1975. She did this because Viet Namese students could make no sense out of English modals, and this helped her a great deal. What I am trying to illustrate here is that one must be careful not to rely too heavily on what one has been taught about the structure of any language, but he must seek, through a study of the use of forms, exactly which meaning can be attributed to which morpheme. Only through this procedure can we ever hope to understand how languages work. The English modals are:

will	intent shall - an allomorph
would	conditional intent
should	strong outside motivation
may	weak possibility
might	conditional possibility
can	ability or permission
could	conditional ability
must	absolute imperative

Note that the last morpheme in the paradigm does not have a conditional because it is absolute. No matter what these forms might have been derived from historically, a serious long-range in-depth study found that these were the meanings which most adequately explained the usages found in the materials studied. Many scholars will want to include "ought to" with the modals, but this clearly is a sequence of morphemes which replaces "should" in certain sequences. Teaching that should, would,

could and might are past tense is absolutely in error, because these are used in future time more than in any other. We have the convention that if the main verb in a sentence is in the past tense, we use these forms rather than the normal forms, i.e., will, can, shall and may for stylistic purposes, but this in no way means that these should be associated with past tense forms of verbs. They operate in the structure of the language in a totally different manner. This is just an example of what I mean by, be sure that you attribute only the meaning to a word that that word carries. If you want to express future time in English, you have no alternative but to give a future time, such as, tomorrow.

Let us return now to the remainder of the sentence given above, i.e., I hope Mary will go home with Bach. The predicate of this sentence clearly is go, a noun-verb which always functions as the predicate when following modals, and the complement of the predicate is home. The phrase with Bach is inherently ambiguous, and the ambiguity can only be resovled by prior knowledge on the part of the listener. If Mary's best friend is Bach, then we will guess (but guess only) that she is riding with her best friend as she goes home. If she has a very special record of some of Bach's music, we may well assume that she is taking this record home with her for some purpose, perhaps to study for an exam in a music class, and so on. Here again, the only point is that we must, as linguists, attribute to the symbolic system only those things which are indeed a part of that system. Much of communication is not langauge, and far too much has been made of attempting to explain ambiguities through language rather than realizing that those ambiguities can only be resolved by additional information which has nothing to do with language or deep structure.

Now we come to the rough part. That is, what goes on inside the head of the speaker. It is, I think, relatively easy to follow the process inside the head of the listener. It is not so simple to

to understand how the thinking process proceeds. Do we think in a deep structure and then convert those thoughts through a set of transformations into surface structure? Clearly, I nor anyone else, actually knows the answer to that. The only thing that I can report which sheds much light on the subject is the fact that when you are attempting to teach English to speakers of other languages, they invariably think in the sequence that the surface structure of their language gives them. Two fundamental things about the nature of real languages makes me believe, and this is a belief, just as Chomsky believes in deep structure, only, that basically all complex thought processes occur in the surface structure of the language one is utilizing at a given moment. The evidence for this belief is the one given above. I have worked with speakers of a number of languages and many of their most difficult problems are related to the fact that they sequence things in the wrong order, and this is not in any "natural" order which could be transformed into English surface structure, but rather it is in the sequence of the surface structure of their native language. If there were a deep structure guiding their thoughts, it seems reasonable to me that one would see foreigners from many countries making all the same sequential mistakes, but my studies have demonstrated to me that this clearly is not the case. When I try to speak in Turkish, one of my most common errors is to think, subject, verb, and then realize that I have forgotten to put in the object. Then I have to go back and begin the sentence over again, in order to put the object in before the verb. The second bit of information about the sequencing of ideas in creating sentences in languages is the fact that if we, indeed, did think in some kind of natural order, this order would show itself in the structure of all languages. This clearly is not the case. I listened to a number of papers at a meeting once wherein scholars had been studying this process and every logical possibility for sequencing of morphemes occurred in various lan-

guages. This implies to me that it is a virtually sure thing that whatever mechanism for thought exists inside the brain, there is no natural sequence which is then "transformed" into surface structure.

Looking at all of the evidence that we have seen about languages of the world, language learning in children and so on, what sort of mechanism would be absolutely necessary to explain both the rigidity and flexibility of language structures? First, if we consider carefully what happened to the children in the second year of their life in the study discussed earlier, we see that the way the language structure developed in each child was radically idiosyncratic. This implies an extremely flexible biological system. In fact, about the only thing that appeared to be shared was an ability to generalize on the basis of two or three experiences and to associate a sequence of noise with some other sensory experiences.

How can this be done? If we think for a moment about what we have seen, we could assume only that the brain has in it someplace a storage compartment which has bipolar storage areas. There might well be no real system to this. It could be a random access storage, and the information stored at one end of the bipolar storage can always be recalled as associated with whatever is stored at the other end. The biological reason for this might have been originally to make it possible for an animal to always associate a certain action in his environment with pain, so that he could avoid that painful experience. If such a mechanism existed, then all that would have to happen for language to develop would be for human beings to have, perhaps, a much larger such storage compartment than other animals, and develop the facility for storing verbal symbols at one end of this device and an array of physical stimulations at the other end. Another assumption would have to be made, too, and that is that the brain has some way of programming the search pattern. We have discussed this briefly in

earlier chapters, but let us see how this might be related to the way languages are learned.

Remember that in all cases children learned single words first, both in our study and other similar studies. At first none of these show any signs of differentiating classes. All words are used as specific labels for classes of phenomena in the real world. Later, the child begins to utilize certain words in first position and certain other words in second position, based on his observance of those around him. Hence, he has established the concept of classes. Perhaps this means that in real terms he has marked off a spot in his brain to place those words that go first and another spot to put those that go in second position. This is precisely the way one programs a computer. Then whatever mechanism controls the brain's functioning goes first to the place where the first class is stored and selects a word. After this is done, it goes to the second spot and selects an appropriate word from that class. As the child grows and his ability to control language increases, the number of classes increases, and the classes are ogranized in the storage compartment of the brain such that a word from those elements which come first are selected first, those that can come after a specific one is selected for first position, and so on down the line. The classes are established in the order that they are learned, and they would be learned more or less in the order of their frequency of use, i.e., those things that are heard very commonly are likely to be learned first, and those things that are not utilized very much are likely to be learned fairly late in the game. Hence the programming occurs because of the nature of the system itself and the way classes are identified by the child. This probably accounts for the fact that different people have, in their brains, different sets of classes and consider different things to be grammatical or ungrammatical. If this type of thing goes on inside the human head when one is grow-

ing up and learning his language, we would, then, expect to see exactly what we see when we look at the surface structures of languages around the world. That is, the only similarities which are bone fide similarities would be very very general sorts of things, such as, all languages have phonological systems which consist of units some of which are crisp (consonants) as opposed to those which are fuzzy (vowels). On a meaningful level, morphemes belong to classes, but the classes from language to language bear little or no relationship to each other, except to the extent that these similarities are a by-product of a method of analysis or description which was utilized very heavily in the nineteenth century. Classes are, indeed, not based on anything in reality, but on things that have occurred in the life-history of the individual. Grammatical categories are the same from person to person, dialect to dialect, language to language and culture to culture exactly to the extent that the life-histories of the individuals which make up the speech communities under consideration are the same. I see no other possible way to explain the vast differences observed in the surface structures of languages around the world.

To begin at the bottom of the scale and summarize our picture of language, we see that there is not a level of analysis which one could call the "distinctive feature" level. Rather, we see that the phonological level consists of a group of consonant or consonant-like units which are identified on the basis of distinctive features and other units, i.e., vowel-like units, which are identified as fuzzy sets on the basis of a number of intersecting facts, e.g., Markov process, likelihood that one or another morpheme will occur in a given situation. Above this level is that of the morpheme, and here again we see that there are fuzzy sets, such as nouns, in some languages, as well as crisp sets, such as pronouns, in other languages. The fundamental differentiation between types of

morphemes in all languages is not that of stems vs affixes, but rather between major and minor morphemes as defined earlier. In general, major morphemes are fuzzy sets and minor morphemes are crisp sets understandable in terms of semantic distinctive features or semantic components, such as, noun formative, first person singular, etc. One problem with understanding languages has been that scholars have attempted to understand all languages on the same basis as they have highly inflective languages, and we have seen in this brief report that isolating languages function quite differently from inflective languages and our theory of language has to be modified to accommodate all types of languages.

Just to consider the concept of deep structure for a moment, when we note that one language expresses the person, number and gender with minor morphemes, another with major morphemes while a third does it both ways, we must conclude that the deep structure has little to do with this in any way at all. Further, even when we say that all languages express person, number and gender, are we really saying anything? Certainly not very much, when we consider that languages all have different numbers of persons, i.e., not necessarily one, two, three, or even singular vs plural, because some languages have nouns to express these ideas and they are pluralized as any other noun is. One language has a valid class noun, while another language does not. One language has a class adjective and another does not. The one single thing that appears to be true for all languages is that they all have classes, they all have two types of classes and these two types of classes appear to work together in a similar manner to express ideas in all languages. Hence any statement of deep structure which is predicated on the existence of classes such as "nouns," "verbs," "animate," "inanimate," etc. is postulating restrictive concepts which are not based on any real data. Even when two languages have an animate-inanimate distinction, what is in the class animate is very dif-

ferent from language to language. This alone should demonstrate beyond any reasonable doubt that the classes are not built on anything in reality. Even the two types of classes appear to have nothing to do with the basic wiring in the human brain. Rather, they probably were built into all languages through history because languages need to be somewhat rigid. If they are too flexible, they will not function as a communicative vehicle. On the other hand if they are too rigid, they will not give communication the flexibility it needs. Hence, it is quite probably that the dichotomy into crisp and fuzzy sets was derived, not from anything wired into the brain, but rather gradually as language systems developed in response to the need for rigidity and flexibility. In the end, we must conclude that what passes now for "deep structure" which is supposed to be wired into the brain of the human animal is reached through intuitive analysis which basically is a reflection of what we have been taught by the Latin and Greek grammarians and the methods of handling linguistic data handed down by them; NOT biology.

Returning now to our summary of language systems, we see that another problem that has come down to us from our Indo-European tradition is the definition of morphology as the "process of word formation." Clearly, this is lumping two levels together. To the extent that a language is purely isolating, it will have only morpheme or only word classes, and could not have both. To the extent that a language is highly inflective, it might well have very few morpheme classes which are radically different from the word classes, if we make the old distinction between stems and affixes, but most languages are somewhat isolating and somewhat inflective, so the traditional dichotomy does not work, and we are forced to look at languages as having a morpheme level of analysis as well as a word level of analysis, because at least some languages have different morpheme classes from word classes.

Above the word level of analysis is that of the phrase, and this level is fairly well understood. In fact, one of the basic reasons for the development of transformational grammar was the fact that many structuralists had concentrated almost exclusively on the phrase level of analysis, and many scholars felt that this left out many very important parts of linguistic structures. This certainly was true, but we accomplish nothing by then going directly to a sentence level of analysis and looking only at sentence formation. Language consists of a series of levels which are hierarchical in nature and which interlock. One might well call this type of view of linguistic structure as *interlocking level* analysis, because I see languages as a number of separate systems all working in harmony together but which are hooked together as cogs in a giant machine. When we erroneously look at two different systems as if they were actually only a single system, we obscure very important differences between languages around the world.

There are types, or classes, of phrases, and a description of this level is an explanation of the different types and a set of rules for combining these types into clauses, at least in English, but it is certainly not clear at this point in time that all languages have the same number and same types of levels. As mentioned earlier, a pure isolating language could not possibly have different word and morpheme classes, hence they would have only a single level for both of these, but there are no pure isolating languages in reality.

Above the phrase, in English at least, we see a clause level, but here again, we need much more careful study of different languages to see if one needs a clause level for all languages, that is, does the clause level, the phrase level and the sentence level all give different classes with different rules for combination to form units on some higher level?

Now we see the nature of the system in toto. Each level consists of units and rules for combining these units to form units on the next higher

level, that is, we have an inventory of phonemes and a set of rules for combining phonemes into morphemes, an inventory of morpheme classes and a set of rules for combining these into words, an inventory of words and a set of rules for combining these into phrases, an inventory of phrases and a set of rules for combining these phrases into clauses, an inventory of clause types and a set of rules for combining these classes of clauses into sentences. At this point it is not at all clear that all languages have the same number of levels, and we know virtually nothing about what levels might possible exist above that of the sentence, but it is clear that something up there does exist.

What happens when we think? We experience sensations, such as the visual images that come to us through our eyes, sounds through our ears, as well as other sensations and we can recall earlier sensations from our memory. The combination of these sensations stimulate our thought processes, but only the very simplest of these are "natural," because virtually all complex thought is in terms of verbal symbols. Even when we think silently to ourselves, much muscle activity goes on as our articulators attempt (or prepare) to formulate the verbal symbols that we are thinking. The programming of this computer-like mechanism has been achieved through the years of learning and building up the classes of units on each level as our language usage increases with age and the process of learning. As soon as we have some sensation, our brain selects some word from those that can come in first position. Once this word has been selected, a definite order must follow, and the brain selects a word from each of the possible classes in the proper sequence to express something related to what he is feeling, or often he simply reacts to a verbal symbol with a verbal symbol without ever reacting to the meaning of the symbol at all. It is perfectly easy for a five-year-old child to create sentences that he has never heard, because he has classified the words, morphemes, phrases, etc., into classes and in formulating a

sentence, he selects the word, morpheme, or whatever, from the proper class with the meaning that fits the situation or sensations that he is having. It is precisely this fact that accounts for the differential treatment of what is grammatical and what is meaningful for different speakers of any language.

We have, of course, not exhausted the work that linguists must do to understand languages or linguistic analysis. What we hve attempted to do in this brief book is point out some of the basic weaknesses in present appoaches to language study and show the direction in which research should go if we are to achieve our ultimate goal, that of understanding not only how languages work but also how people's minds work as they attempt to speak to each other. I trust that the book will stimulate a lot of controversy and that this controversy will stimulate research which will in the end develop accurate and reliable descriptions of the actions of human beings when they react to each in various cultures around the world.

SELECTED BIBLIOGRAPHY

Bach, Emmon; *An Introduction to Transformational Grammars;* Holt, Rinehart and Winston, 1964.

Bloomfield, Leonard; *Language;* Holt, Rinehart and Winston, 1933.

Chomsky, N.; *Syntactic Structures;* Mouton, 1957.

Fries, C. C.; *The Structure of English. An Introduction to the Construction of English Sentences;* Harcourt, Brace and Company, 1952.

Garvin, Paul L.; (a review of) *Preliminaries to Speech Analysis;* Language, Volume 29, Number 4, pp. 472-481, 1953.

Greenburg, Joseph H.; *A Quantitative Approach to the Morphological Typology of Language;* IJAL, Volume 26, pp. 178-194, 1960.

Hockett, Charles; *A Manual of Phonology;* Indiana University Press, 1955.

Jakobson, Roman with Gunnar M. Fant and Morris Halle; *Preliminaries to Speech Analysis;* MIT Research Report Number 13, 1952.

Jakobson, Roman with Morris Halle; *Fundamentals of Language;* Mouton, 1956.

Joos, Martin; *Acoustic Phonetics;* Language Supplement, 1948.

Kroeber, A. L.; *Culture; the Superorganic;* in *The Nature of Culture;* University of Chicago Press, pp. 22-50, 1952.

Pierce, Joe E.; *Spectrographic Study of Vocalic Nuclei;* Language Learning, Volume XII, Number 3, pp. 241-247, 1963.

Phonemic Theory and the Analysis of English Syllabic Nuclei; Linguistics, Volume 7, pp. 63-82, 1964.

The Phonemic Composition of English Morphemes; Linguistics, Volume 20, pp. 60-71, 1966.

The Morphemes of English: Major Morpheme Stem Classes; Linguistics, Volume 36, pp. 29-51, 1967.

A Look at the So-called -ing-forms of English Verbs; Linguistics, Volume 50, pp. 59-69, 1970.

A Spectrographic Analysis of English Vowel Phonemes under Primary Stress: Linguistics, Volume 84, pp. 41-84, 1972.

Life Histories of Individuals and their Impact on International Communications; in *International Communications;* Fred L. Casmir (ed.), University Press of America, pp. 522-540, 1978.

How English Really Works; The HaPi Press, 1979.

Pierce, Joe E. with Inga Vanwaardenburg Hanna; *The Development of a Phonological System in English Speaking American Children;* The HaPi Press, 1974.

The Development of a Linguistic System in English Speaking American Children; The HaPi Press, 1978.

Pike, Kenneth L.; *Phonemics;* University of Michigan Press, 1947.

Sapir, Edward; *Language;* Harcourt, Brace and World, Inc., 1921.

Selected Writings of Edward Sapir in Language, Culture and Personality; David G. Mandelbaum (ed.), University of California Press, 1949.

Saussure, Ferdinand de; *A Course in General Linguistics;* translated by Wade Baskin, Philosophical Library, New York, 1959.

Smith, Henry Lee; (Films on Linguistics) *Languages and Linguistics;* SUNY Buffalo, circa 1960.

Southworth, Franklin C. with Doswani, C. J.; *Foundations of Linguistics;* The Free Press, 1974.

Trager, George L with Henry Lee Smith; *The Outline of English Structure;* Studies in Linguistics, Occasional Papers, Number 13, Battenburg Press, 1951.

Trubetskoy, N. S.: *Grundzuge der Phonologie;* Cercle Linguistic de Prague, 1939.

Wells, Rulon S.; *Saussure's System of Linguistics;* Word, Volume 3, Number 1, pp. 15-18, 1947.

Whorf, Benjamin Lee; *Language, Thought and Reality. Selected Writings of Benjamin Lee Whorf;* John B. Carroll (ed.), Cambridge: Technology Press, 1956.

Zadeh, L. A.; *Fuzzy Languages and their Relation to Machine Intelligence;* Electronic Research Laboratories, University of California Press, ERL-M302, 1971.

INDEX

Accusative, 66, 67
Acoustic Correlates, 9-12, 18, 32-5, 137
Acoustic Phonetics, 10, 15-8, 31-5, 43-4, 137
Acoustic Spectrograph, 10, 17, 21, 34-5
Adjectives, 62-3, 67
Adverbs, 62-4, 67, 115
Affixes, 50, 52-3, 66, 69-70, 81-8, 110-3, 152-3
Agglutinative Languages, 49, 52-3, 69, 73, 110,113
Algonkin Languages. 127
Allophones, Positional, 34-5, 40
Ambiguity, 141-7
Animate-Inanimate Nouns, 127
Anthropological Linguistics, 1
Anthropology, 4-6
Articles, 82, 121
Articulation, 10-2, 17-8, 34-5, 94-5
Aspiration, 11
Bach, Emmon, 14, 54-5
Bloomfield, Leonard, 24-5
Bostain, 113
Bui Thi Hong Nguyet, 246
Chomsky, N., 3, 55, 75, 122-4, 148
Clauses, 105, 128-9, 154-155
Closed Corpus, 20-3
Cloze Procedure, 37
Communication, 39-40, 82, Chapter VI, 138, 147, 153
Competence, 39-40
Complements (verbal), 77, 80, 129, 141, 147
Composite Verbs, 115
Computer, 7, 42
Conditioning, 54, 95-100
Conjunctions, 82-3
Content Words, 57
Creativity, 104-5
Culture, 4-6
Deep Structure, 2-3, 6, 35-9, 46-7, 70-2, 82, 90-1, 98-101, 107, 109, 122-6, 133-6, 140-2, 147-55
Derivation, 63, 82-9, 111, 113
Descriptive Linguistics, 19, 133
Dialects, 2
Diffusion, 4
Discovery Procedures, 20-2, 54, 64, 67, 69, 133
Distinctive Features, 2, 9-7, 25-37, 43, 47, 90, 226, 130, 134, 136, 151, 152
Doswani, C. J., 19
English, 2, 10-5, 22-3, 29, 33-7, 40-3, 47, 50-60, 80, Chapts. III, IV, V,

English (continued), 133, 136-7, 143-6
Extrapolation, 97-8
Fant, Gunnar M., 9, 13-4, 26, 28
Fries, Charles C., 50-1, 57
Function Words, 51
Garvin, Paul L., 17, 28
Generative Grammar, 20
Gerunds, 77-8, 84-6, 88, 111, 121
Grammar, 25, 38-9, Chapt. III, 96, 102-3
Greek Grammarians, 24, 49-50
Greenberg, Joseph H. 53
Halle, Morris, 9, 13-4, 26, 28
Hanna, Inga Vanwaardenburg, 91
Haskins Laboratories, 14-6, 44
Hockett, C., 19-28
Imbedding, 145
Immediate Constituent Analysis, 142-4
Infinitives, 116, 121
Infixes, 50
Inflection, 52, 62-3, 67, 82-9, 120
Inflective Languages, 4, 9, 50-4, 63, 69, 73, 81, 110, 113, 152-3
Innate Grammar, 3-4
Inorganic, 4-9
Interlocking Level Analysis, 154
Intuition, 54-5, 64, 67, 54-5, 64, 67, 72, 79, 127, 153
Isolating Languages, 4, 9, 51-3, 69, 73, 80-1, 88, 90, 110, 113, 115, 118, 152-3
Jakobson, Roman, 9, 12-4, 16, 26, 28, 35, 92, 130
Japanese, 38, 55-6, 71-4, 81, 106-7, 123
Joos, Martin, 10, 15, 17, 32-3, 37
Kernel Sentences, 110, 118
Kinesics, 27, 141
Kroeber, 4-6
Language Acquisition, 66, 72, 76, Chapter IV, 149-51
Language-Culture, 106-8
Langue, 17-8
Life Histories, 27, 39, 45, 47, 126, 131, 141, 144, 151
Linguistic Universals, 15, 124
Logic, 5-9, 12, 18, 25, 34
Major Morphemes, 52-3, 57, 64-8, 71-8, 81-3, 90, 110, 118, 152
Markov Process, 17, 35-8, 40, 43, 45, 47, 116, 133-4, 137-45, 151
Maturation, 93, 95-6, 100
Meaning, 26-7, 30-1, 29, 62, 77, 80-4, 98, 114, 127, 140-7, 151, 155
Minor Morphemes, 52-3, 57, 64-9, 71-83, 90, 110, 113, 118-9, 152
Modals, 145-7

Morphology, 42-7, Chapt. III, 116, 118, 120, 124-31, 140-44, 151-55
Nature, 5-7
Nominal Function, 121
Nouns, 19, 46, 54-62, Chapt. III, 119-20, 127, 151
Noun-Verbs, 59-62, 68, 77, 83, 86-8, 120, 121
Organic, 5-9
Ought to, 146
Paralanguage, 27, 141
Parole, 17-8
Participles, 82-8, 121
Particles, 67, 82, 111
Parts of Speech, 51
Passive Voice, 120, 142-3
Past Tense, 23, 120
Perception, 2, 98-100
Phonemic Overlap, 34, 36, 40
Phonology, 9-18, 24-39, Chapt. II, 82, 90-104, 116, 129-40, 144, 151, 154
Phrases, 109-19, 124, 126-8, 144, 154-5
Pierce, Joe E., 12, 14-4, 20, 42, 91, 97, 108
Pike, Kenneth, 12-3, 16
Predicate Function, 86-7, 115, 120-1, 127-8, 144-5, 147
Prefixes, 50
Prepositions, 82, 113, 119
Processing Language in the Brain, Chapt. VI
Pronouns, 55-6, 69-75, 81-2, 133, 145, 151
Question Words, 78-82
Redundance, 27
Reinforcement, 94-5, 98-100
Sample Selection, 91-3
Sapir, Edward, 1, 24-5, 49, 106
Saussure, E., 17, 24
Science, 5-12
Semantic Components, 57, 73, 75-6, 82, 152
Sentences, 109, 111, 127-9
Sets, Crisp, 7-18, 43-8, 73, 75, 82-3, 116, 130-1, 151-3
Sets, Fuzzy, 7-18, 40, 43, 45, 47-8, 57, 61-2, 68, 73, 82-3, 90, 116, 130-1, 139, 151-3
Smith, Henry Lee, 50
Social Context, 41, 43-5, 70, 141
Southworth, Franklin C., 19
Speech Community, 18
Statistics, 18, 61, 125-6
Stems, 50, 52-3, 57, 66, 88, 152-3
Stockwell, Robert, 15
Structuralism, 1-3, 20-1, 24-5, 34, 143, 154
Sturtevant, E. H., 18
Subject-verb inversion, 79
Suffixes, 50
Superorganic, 4-10
Surface Structure, 100, 122-3, 148-51

Syllables, 38
Syntax, 25-6, Chapt. V, 131, 154-5
Synthetic Languages, 45, 52, 73
Thinking, 147-9, 155-6
To be, 76-9, 85-7, 121, 145
To do, 78-9
To have, 78-9
Trager, George L., 130
Transformational Grammar, 20-1, 25-6, 44-5, 123-5, 141, 144, 148, 154
Trubetskoy, N. S., 24
Turkish, 2, 34, 59, 65-75, 83-4, 106-7, 110, 113, 123, 148
Tylor, E., 6
Typology, 49-57
Verbs, 21-3, 54, 58-62, Chapt. III, 120
Verbs, Irregular, 78, 87
Voegelin, Carl F., 52
Wells, Rulon S., 17
Whorf, Benjamin Lee, 106
Words, 83-102, 111-2, 117-8, 120, 124-30, 140, 144, 154-5
Zadeh, L. A., 7, 13, 18

Other Hapi Press Books

Mayah Balse; Mystics and Men of Miracles in India, $8.50.

Joe E. Pierce; A Linguistic Method of Teaching Second Languages, 6.95.

Egbert S. Oliver; The Tarbells of Yankton, $8.95.

Richard L. Brinkman; Cultural Economics, $15.95.

Abdul Qayum; Social Cost-Benefit Analysis, $9.95.

Joe E. Pierce; How English Really Works, $9.95.

Egbert S. Oliver; The Shaping of a Family, $8.95.

Joe E. Pierce; A Theory of Language, Culture and Human Behavior, $7.95.

Ethne K. Marenco; The Transformation of Sikh Society, $8.95.

Gwendolyn Pierce; An Exploration into the Applicability of a Psychological Technique for Anthropological Research, $2.50.

Joe E. Pierce and Ingrid Hanna; The Development of a Linguistic System in English Speaking American Children, $7.95.

Joe E. Pierce and Ingrid Hanna; The Development of a Phonological System in English Speaking American Children, $6.95.

B. Stefaniszyn; African Lyric Poetry, $5.95.

Joe E. Pierce; The Sapien Homo, $7.95.

Joe E. Pierce; The Bitter Winds, $7.95.